

(19)대한민국특허청(KR)

(12) 공개특허공보(A)

(51) Int. Cl.⁷
H04N 7/30

(11) 공개번호 10-2004-0004178
(43) 공개일자 2004년01월13일

(21) 출원번호 10-2003-0045330
(22) 출원일자 2003년07월04일

(30) 우선권주장 60/393,894 2002년07월05일 미국(US)

(71) 출원인 마이크로소프트 코퍼레이션
미국 워싱턴주 (우편번호 : 98052) 레드몬드 원 마이크로소프트 웨이

(72) 발명자 토우라피스, 알렉산드로스
사이프러스니코시아트세리2480에스터.일라리오나스5

리,시펑
미국92618캘리포니아주어바인오클렌628

후,펑
중국100080베이징빌딩피프쓰원싸우전드투핸드레드나인쓰하이드리안디스트릭트즈히진파젠
다

술리반,가리
미국98053워싱턴주레드몬드투핸드레드투엔티세븐쓰애비뉴엔이6239

(74) 대리인 주성민
백만기
이중희

심사청구 : 없음

(54) 데이터 압축을 위한 개선된 최적화 기술

요약

데이터, 예를 들어, 이미지, 비디오 등을 인코딩하는 것에 관한 방법 및 장치가 제공된다. 예를 들어, 어떤 방법은, 복수개의 서로 다른 양자화 함수를 사용하여 복수개의 대응하여 양자화된 데이터 부분을 생성하기 위해서 데이터의 적어도 일부분을 처리하는 단계, 및 적어도 하나의 임계치에 기초하여 상기 양자화된 데이터 부분들 중 하나를 선택적으로 출력하는 단계를 포함한다. 또한, 그 방법은 초기 데이터를 복수개의 부분으로 분할하는 단계 및 적어도 하나의 분류 특성에 기초하여 데이터 부분을 분류하는 단계를 포함한다. 여기서, 예를 들어, 분류 특성에 관련된 임계치가 있을 수 있다. 추가적인 구문이, 여러 다른 모션 예측 케이스(motion prediction case)를 허용함으로써 상당히 높은 압축 효율이 가능하도록 채택될 수 있다. 정적 시간 예측기(stationary temporal predictor) 뿐만 아니라 공간 모션 벡터 예측(spatial motion vector prediction)도 고려하는 고효율의 타임 스탬프 독립 다이렉트 모드(time stamp independent Direct Mode)가 제공된다.

대표도

도 2a

색인어

데드존 양자화기, 데이터 압축, 레이트 디스토션 최적화

명세서

도면의 간단한 설명

도 1은 본 발명의 일 구현예에 따라 사용하기에 적합한 컴퓨터 형태의 예시적인 장치를 도시하는 블록도.

도 2a 내지 도 2d는 본 발명의 일 구현예에 따라 예시적인 선택가능 데드존 양자화기를 도시하는 그래프.

도 3은 본 발명의 일 구현예에 따라 서로 다른 양자화 프로세스를 데이터에 선택적으로 적용하기 위한 예시적인 방법을 도시하는 흐름도.

도 4는 본 발명의 일 구현예에 따라 서로 다른 양자화 프로세스를 데이터에 선택적으로 적용하기 위한 예시적인 로직을 도시하는 블록도.

도 5는 본 발명의 다른 구현예에 따라 서로 다른 양자화 프로세스를 데이터에 선택적으로 적용하기 위한 예시적인 로직을 도시하는 블록도.

도 6a 내지 도 6d는 본 발명의 또다른 구현예에 따라 서로 다른 양자화 프로세스를 데이터에 선택적으로 적용하기 위한 예시적인 로직을 도시하는 블록도.

도 7은 본 발명의 일 구현예에 따라 서로 다른 양자화 프로세스를 데이터에 선택적으로 적용하기 위한 예시적인 로직을 도시하는 블록도.

도 8은 본 발명의 일 구현예에 따라 이동 벡터에 대한 서로 다른 예측 방법을 선택적으로 적용하기 위한 로직에 사용되는 예시적인 구문 정보를 나열하는 차트.

도 9는 본 발명의 일 구현예에 따라 서로 다른 인코딩 방식을 비디오 데이터에 선택적으로 적용하기 위한 예시적인 로직을 도시하는 블록도.

도 10은 본 발명의 다른 구현예에 따라 서로 다른 인코딩 방식을 비디오 데이터에 선택적으로 적용하기 위한 예시적인 로직을 도시하는 블록도.

도 11은 본 발명의 일 구현예에 따라 선택적으로 적용된 인코딩 방식을 채용하여 비디오 시퀀스의 어떤 특징을 도시하는 도면.

도 12는 본 발명의 일 구현예에 따라 다이렉트 모드 비디오 시퀀스에 대한 시공간 예측을 위한 예시적인 방법을 도시하는 흐름도.

<도면의 주요 부분에 대한 부호의 설명>

120: 컴퓨팅 환경 132: 처리부

134: 시스템 메모리 136: 버스

148: 자기 디스크 152: 광디스크

158: 운영 체제 160: 애플리케이션 프로그램

166: 키보드 172: 모니터

202, 204, 206, 208: 데드존 양자화기

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 컴퓨터 및 유사 장치에 관한 것으로, 특히, 데이터를 압축/인코딩하고 압축풀기/디코딩하기 위한 방법, 장치 및 시스템에 관한 것이다.

데이터, 특히, 이미지 데이터와 비디오 데이터, 압축/인코딩하고 압축풀기/디코딩하기 위한 개선된 방법 및 장치에 대한 지속적인 요구가 있어 왔다. 코딩 효율의 개선은 보다 많은 정보가 컴퓨터 및 다른 유사 장치에 의해 보다 용이하게 처리, 전송 및/또는 저장될 수 있게 한다. 인터넷과 다른 유사 컴퓨터 네트워크, 및 무선 통신 시스템의 수적 증가에 기인하여, 가용 자원을 충분히 활용할 고효율의 인코딩 기술을 제공하는 것이 요구된다.

레이트 디스토션 최적화(Rate Distoriton Optimization; RDO) 기술은, 종래의 인코딩 방식에 비해 인코딩 효율을 상당히 개선할 수 있기 때문에, 비디오 및 이미지 인코딩/디코딩 시스템에 매우 광범위하게 사용된다.

추가적인 정보는, 예를 들어, Trac Duy Tran 이 1994년 5월에 Massachusetts Institutue of Techonology 에서 'A Locally Adaptive Perceptual Masking Threshold Model for Image Coding'이라는 제목의 컴퓨터 과학기술의 석사학위 논문에서 발견될 수 있다.

보다 높은 인코딩 효율을 제공하고자 하는 지속적인 욕구가 있기 때문에, 기존 시스템에 비해 개선된 코딩 효율을 달성하기 위해 ROD 또는 다른 유사 기술의 성능을 더욱 향상시키는 개선된 방법 및 장치가 요구된다.

발명이 이루고자 하는 기술적 과제

본 발명은 이미지, 비디오 및/또는 오디오 데이터를 포함하고 이에 한정되지 않는 다양한 형태의 데이터를 압축/인코딩, 압축풀기/디코딩, 및/또는 다른 처리에서 사용될 수 있는 개선된 방법 및 장치를 제공한다.

상술한 요구 및 다른 요구들은, 예를 들어, 복수개의 서로 다른 양자화 함수(quantization function)를 사용하여 적어도 데이터 부분을 처리하여 복수개의 대응 양자화 데이터 부분을 생성하는 단계와 적어도 하나의 임계치에 기초하여 양자화된 데이터 부분들 중의 하나를 선택적으로 출력하는 단계를 포함하는 방법에 의해 충족된다. 또한, 그 방법은 초기 데이터를 복수개의 부분으로 분할하는 단계와 적어도 하나의 분류 특성(classification charateristic)에 기초하여 그 데이터 부분을 분류하는 단계를 포함할 수 있다. 이 때, 예를 들어, 그 분류 특성에 관련되는 임계치가 있을 수 있다.

예를 들어, 초기 데이터는 이미지 데이터, 비디오 데이터, 오디오 데이터, 스피치 데이터(speech data) 등을 포함할 수 있다. 선택되는 부분은 블록, 매크로블록(macroblock), 슬릿(slit), 슬라이스(slice), 섹션(section) 등의 형태를 취할 수 있다. 그 분류 특성(들)은 에지(edge) 특성, 텍스처(texture) 특성, 평활(smoothness) 특성, 루미넌스(luminance) 특성, 크로미넌스(crominance) 특성, 컬러 특성, 노이즈 특성, 오브젝트 특성, 모션(motion) 특성, 사용자 선호도 특성, 사용자 인터페이스 초점(user interface focus) 특성, 레이어링(layering) 특성, 타이밍 특성, 볼륨(volume) 특성, 주파수 특성, 피치(pitch) 특성, 톤(tone) 특성, 품질 특성, 비트 레이트(bit rate) 특성, 데이터 유형 특성, 해상도(resolution) 특성, 암호화 특성등을 포함할 수 있다.

예시적인 일 구현예에서, 복수개의 서로 다른 함수는 2개 이상의 서로 달리 동작하는 데드존 양자화기(Deadzone Quantizer)를 포함할 수 있다. 이 때, 예를 들어, 데드존 양자화기는, 어댑티브 커버리지(adaptive coverage) 데드존 양자화기, 가변 커버리지 크기(variable coverage size) 데드존 양자화기 동일 수 있다. 또한, 그 방법은 양자화 부분을 인코딩하는 단계를 포함할 수 있다. 그 방법은 양자화된 데이터 부분을 선택하기 위해 레이트 디스토션 최적화(RDO; rate distortion optimization)를 이행하는 단계를 더 포함할 수 있다.

본 발명의 또다른 예시적인 구현예에서, 다른 방법은, 이미지 데이터의 적어도 일부분에 대하여 적어도 하나의 특성 분석(characteristic analysis)을 이행하는 단계와 그 특성 분석에 기초하여 인코더 내에 적어도 하나의 어댑티브 양자화 파라미터(adaptive quantization parameter)를 선택적으로 선택하는 단계, 및 그 인코더를 사용하여 이미지 데이터 부분을 인코딩하는 단계를 포함한다.

또다른 예시적인 구현예에서, 적어도 2개의 서로 다른 데드존 양자화기를 이용하여 이미지 데이터의 적어도 일부분이 인코딩되게 하는 단계, 및 적어도 하나의 판정 팩터(decision factor)에 관련되는 레이트 디스토션 최적화(RDO)에 기초하여 적어도 2개의 서로 다른 데드존 양자화기의 출력 내 바람직한 인코드 데이터를 식별하는 단계를 포함하는 방법이 제공된다.

또한, 상술한 요구 및 다른 요구들은, 제1 데드존 양자화기를 이용하여 이미지 데이터의 적어도 일부분을 인코딩하는 단계, 제1 데드존 양자화기가 적어도 하나의 판정 팩터를 충족하는 지를 판정하는 단계, 및 만약 그러하다면, 제1 데드존 양자화기의 출력을 출력하는 단계를 포함하는 방법에 의해 충족된다. 만약 그렇지 않으면, 제1 데드존 양자화와 다른 적어도 제2 데드존 양자화기를 이용하여 이미지 데이터의 부분을 인코딩되게 한다. 이 때, 예를 들어, 그 방법은 RDO 판정 또는 다른 판정에 기초하여 이미지 데이터 부분의 허용가능한 인코딩된 버전을 식별하는 단계를 더 포함할 수 있다.

또다른 예시적인 구현예에서, 이미지 데이터의 적어도 일부분에 대하여 이미지 분석을 이행하는 단계, 분석된 이미지 데이터 부분에 대하여 차단 분류(block classification)를 이행하는 단계, 차단 분류된 이미지 데이터 부분의 데드존 양자화를 이행하는 단계, 및 데드존 양자화된 이미지 데이터 부분의 인코딩을 이행하는 단계를 포함하는 방법이 제공된다. 이 때, 예를 들어, 그 이미지 분석은 에지 검출 분석(edge detection analysis), 텍스처 분석(texture analysis) 등을 포함할 수 있다.

또다른 예시적인 구현예에 따라, 적어도 2개의 서로 다른 인코더를 이용하여 비디오 데이터의 적어도 일부분이 인코딩되게 하는 단계 -상기 적어도 2개의 서로 다른 인코더는 비디오 이미지 데이터의 논 레지듀얼 모드(Non Residual Mode)를 지원하도록 동작 구성된 데드존 양자화기를 포함함-, 및 적어도 하나의 판정 팩터에 관련되는 레이트 디스토션 최적화(RDO)에 기초하여 2개의 서로 다른 인코더들 중의 하나의 출력 내에 바람직한 인코드 프레임 데이터를 식별하는 단계를 포함한다.

또다른 예시적인 방법은, 이미지 데이터의 적어도 일부분의 적어도 하나의 특성에 기초하여 양자화 함수를 갖는 인코딩 로직을 동작 구성하는 적어도 하나의 라그랑지 곱수(Lagrangian multiplier)를 선택적으로 변경하는 단계, 및 그 인코딩 로직을 이용하여 이미지 데이터 부분을 인코딩하는 단계를 포함한다.

또다른 예시적인 구현에서, 예시적인 방법은 인코더 로직을 이용하여 비디오 이미지 데이터의 적어도 일부분에서 인코딩하는 단계, 및 인코더 로직이 그 인코더 로직에 의해 채용되는 이동 벡터 예측 유형을 식별하는 구문 정보를 출력하게 하는 단계를 포함한다.

비디오 인코딩 관련 정보를 전달하는 데 이용되는 방법은, 비디오 데이터를 인코딩하는 단계, 및 인코드 비디오 데이터 구문 부분 내에 적어도 하나의 디스크립터-상기 디스크립터는 비디오 데이터로 인코딩된 적어도 하나의 B 프레임에 관련되는 어댑티브 공간/시간 인코딩(adaptive spatial/spatio-temporal encoding)을 식별함-를 선택적으로 설정하는 단계를 포함한다. 비디오 인코딩 관련 정보를 전달하는 데 이용되는 다른 방법은, 비디오 데이터를 인코딩하는 단계, 인코드된 비디오 데이터 구문 부분 내에 적어도 하나의 디스크립터(descriptor) -상기 디스크립터는 비디오 데이터로 인코드된 적어도 하나의 인터 프레임(Inter frame) 내에서 어댑티브 카피/모션 카피 스킵 모드(adaptive copy/motion-copy skip mode)를 식별함-를 설정하는 단계를 포함한다.

비디오 프레임의 정적 시간/공간 부분을 고려하여 비디오의 타임 스탬프 독립 모드(time stamp independent mode) 인코딩에서 이용하기 위한 예시적인 방법이 제공된다. 이 때, 예를 들어, 그 방법은 비디오 시퀀스 내에 비디오 프레임의 적어도 일부분에 관련되는 모션의 공간적 예측을 선택적으로 적용하는 단계와, 만약 다른 비디오 프레임의 참조 부분에 대한 시간적 모션 예측 정보가 0이면, 모션의 공간적 예측을 0으로 설정하는 단계를 포함한다.

발명의 구성 및 작용

도면을 참조하면, 유사 참조부호는 유사 구성요소를 나타내며, 본 발명은 적합한 컴퓨팅 환경에서 구현된 것으로서 기술된다. 비록 요구되는 것은 아니지만, 본 발명은, 개인용 컴퓨터의 형태를 취할 수 있는 서버 컴퓨터, 워크스테이션, 전용 서버, 복수개의 프로세서, 메인프레임 컴퓨터 등에 의해 실행되는 프로그램 모듈 등의 컴퓨터 실행가능 명령의 일반 구문에서 기재될 것이다. 일반적으로, 프로그램 모듈은, 특정 태스크를 이행하거나 특정 추상 데이터형(abstract data type)을 구현하는 루틴, 프로그램, 오브젝트, 컴포넌트, 데이터 구조 등을 포함한다. 본 발명은 통신 네트워크를 통하여 연결되는 원격 프로세싱 장치에 의해 태스크가 이행되는 분산 컴퓨팅 환경 내에서 실시될 수 있다. 분산 컴퓨팅 환경에서, 프로그램 모듈은 로컬 및 원격 메모리 스토리지 장치에 배치될 수 있다.

예시적인 컴퓨팅 환경:

도 1은 적절한 컴퓨팅 환경(120)의 예를 도시하며, 그 컴퓨터 환경 상에서 이하 기술되는 방법 및 배치가 구현된다.

예시적인 컴퓨팅 환경(120)은 적합한 컴퓨팅 환경(120)의 일 실시예이며, 여기서 기재되는 개선된 방법 및 배치의 사용 또는 기능의 범주에 대한 어떠한 제한을 암시하려는 것은 아니다. 또한, 컴퓨팅 환경(120)도, 그 컴퓨팅 환경(120) 내에 도시되는 구성요소들 중의 임의의 단독 또는 조합에 관련하여 임의의 종속성 또는 요건을 갖는 것으로서 이해되어져서는 안된다.

여기서, 개선된 방법 및 배치는 수많은 다른 범용 또는 특정 목적 컴퓨팅 시스템 환경 또는 구성을 사용하여 동작한다.

도 1에 도시된 바와 같이, 컴퓨팅 환경(120)은 컴퓨터(130) 형태의 범용 컴퓨팅 장치를 포함한다. 컴퓨터의 구성요소(130)는 적어도 하나의 프로세서 또는 처리부(132), 시스템 메모리(134), 및 시스템 메모리(134)를 포함하는 다양한 시스템 구성요소를 프로세서(132)에 결합하는 버스(136)를 포함할 수 있다.

버스(136)는, 메모리 버스 또는 메모리 컨트롤러, 병렬 버스, 가속 그래픽 버스(accelerated graphics port), 임의의 다양한 버스 구조를 이용하는 프로세서 또는 로컬버스를 포함하는 임의의 여러 유형의 버스 구조들 중의 하나 이상의 버스 구조를 나타낸다. 예를 들어-한정이 아님- 이러한 구조는 산업 표준 구조 (ISA; Industry Standard Architecture) 버스, 마이크로 채널 구조 (MCA; Micro Channel Architecture) 버스, 향상된 ISA (EISA; Enhanced ISA) 버스, 비디오 전자표준협회(VESA; Video Electronics Standards Association) 로컬 버스, 및 메자닌 버스(Mezzanine bus)로 잘 알려진 주변 구성요소 접속(PCI; Peripheral Component Interconnect) 버스를 포함한다.

통상, 컴퓨터(130)는 다양한 컴퓨터 판독가능 매체를 포함한다. 이러한 매체는 컴퓨터(130)에 의해 액세스가능한 임의의 가용매체일 수 있으며, 휘발성 및 비휘발성 매체, 분리형 및 비분리형 매체를 포함한다.

도 1에서, 시스템 메모리(134)는 랜덤 액세스 메모리(RAM; 140)와 같은 휘발성 메모리 형태, 및/또는 판독전용 메모리(ROM; 138)와 같은 비휘발성 메모리 매체의 형태의 컴퓨터 판독가능 매체를 포함한다. 기본 입출력 시스템(BIOS; 142)은, 초기 시동(start-up) 동안과 같이, 컴퓨터(130) 내의 요소들 간의 정보를 전송할 수 있게 하는 기본 루틴을 포함하여, ROM(138)에 저장된다. 통상, RAM(140)은 프로세서(132)에 즉시 액세스 가능하고/가능하거나 프로세서(132)에 의해 현재 실행중인 데이터 및/또는 프로그램 모듈을 포함한다.

컴퓨터(130)는 다른 분리형/비분리형, 휘발성/비휘발성 컴퓨터 스토리지 매체를 더 포함할 수 있다. 예를 들어, 도 1은, 비분리형, 비휘발성 자기 매체(도시 생략, 통상 '하드 드라이브'라 지칭됨)로부터 판독하고 이에 기입하기 위한 하드 디스크 드라이브(144), 분리형, 비휘발성 자기 디스크(148; 예를 들어, '플로피 디스크')로부터 판독하고 이에 기입하는 자기 디스크 드라이브, CD-ROM, CD-R, CD-RW, DVD-ROM, DVD-RAM과 같은 분리형, 비휘발성 광 디스크(152)로부터 판독하고 이에 기입하기 위한 광 디스크 드라이브(150)를 도시한다. 하드 디스크 드라이브(144), 자기 디스크 드라이브(146) 및 광 디스크 드라이브(150)는 하나 이상의 인터페이스(154)에 의해 각각 버스(136)에 접속된다. 하드 디스크 드라이브(144), 자기 디스크 드라이브(146) 및 광디스크 드라이브(150)는 하나 이상의 인터페이스(154)에 의해 각각 버스(136)에 접속된다.

그 드라이브들과 관련 컴퓨터 판독가능 매체는 컴퓨터 판독가능 명령, 데이터 구조, 프로그램 모듈, 및 컴퓨터(130)용 다른 데이터의 비휘발성 스토리지를 제공한다. 비록, 여기서 기재된 예시적인 환경은, 하드 디스크, 분리형 자기 디스크(148) 및 분리형 광디스크(152)를 채용하지만, 자기 카세트, 플래시 메모리 카드, 디지털 비디오 디스크, 랜덤 액세스 메모리(RAM), 판독전용 메모리(ROM) 등과 같은, 컴퓨터에 의해 액세스 가능한 데이터를 저장할 수 있는 다른 유형의 컴퓨터 판독가능 매체가 예시적인 운용 환경에서 또한 이용될 수 있음이 당업자에 의해 인식되어져야 한다.

다수의 프로그램 모듈은 하드 디스크, 자기 디스크(148), 광디스크(152), ROM(138), 또는 RAM(140) 상에, 예를 들어, 운용 시스템(158), 하나 이상의 애플리케이션 프로그램(160), 다른 프로그램 모듈(162), 및 프로그램 데이터를 포함하여, 저장될 수 있다.

여기서 기재되는 개선된 방법 및 배치는 운용 시스템(158), 하나 이상의 애플리케이션 프로그램(160), 다른 프로그램 모듈(162), 및/또는 프로그램 데이터(164) 내에서 구현될 수 있다.

사용자는 키보드(166) 및 포인팅 장치(168; 예를 들어, 마우스 등)와 같은 입력 장치를 통해 컴퓨터(130)에 명령과 정보를 제공한다. 다른 입력 장치들(도시 생략)은 마이크로폰, 조이스틱, 게임 패드, 위성 접시, 직렬 포트, 스캐너, 카메라 등을 포함할 수 있다. 이들 및 다른 입력장치는 버스(136)에 결합되는 사용자 입력 인터페이스(170)를 통하여 처리부(132)에 접속되지만, 병렬 포트, 게임 포트, 또는 공통 직렬버스(USB)와 같은 다른 인터페이스와 버스 구조에 의해 접속될 수 있다.

모니터(172) 또는 다른 유형의 디스플레이 장치는 또한 비디오 어댑터(174)와 같은 인터페이스를 통하여 버스(136)

에 접속된다. 모니터(172)에 더하여, 통상, 컴퓨터는, 스피커와 프린터와 같은 다른 병렬 출력 장치(도시 생략)를 포함하며, 이들은 다른 출력 병렬 인터페이스(175)를 통하여 접속될 수도 있다.

컴퓨터(130)는 원격 컴퓨터(182)와 같은 적어도 하나의 원격 컴퓨터로의 논리 접속(logical connection)을 이용하여 네트워크 환경에서 작동할 수 있다. 원격 컴퓨터(182)는 컴퓨터(130)에 관련하여 여기서 기재되는 요소 및 특징들의 다수 또는 모두를 포함할 수 있다.

도 1에 도시된 논리 접속은 근거리 네트워크(LAN; 177) 및 일반 광역 네트워크(WAN; 179)이다. 이러한 네트워킹 환경은 사무실, 사내 광역 컴퓨터망(enterprise-wide computer networks), 인트라넷, 및 인터넷에서 흔한 것들이다.

LAN 네트워킹 환경에서 사용될 때, 컴퓨터(130)는 네트워크 인터페이스 또는 어댑터(186)를 통하여 LAN(177)에 접속된다. WAN 네트워킹 환경에서 사용될 때, 통상, 컴퓨터는 모뎀(178) 또는 WAN(179) 상에 통신을 확립하기 위한 다른 수단을 포함한다. MODEM(178) -내장형 또는 외장형일 수 있음- 은 사용자 입력 인터페이스(170) 또는 다른 적절한 메커니즘에 의해 시스템 버스(136)에 접속될 수 있다.

도 1에서는 인터넷을 통한 WAN의 특정 구현예가 도시된다. 이 때, 컴퓨터(130)는 인터넷(180)을 통하여 적어도 하나의 원격 컴퓨터(182)와 통신을 확립하기 위해 모뎀(178)을 채용한다.

네트워크 환경에서, 컴퓨터에 관련하여 도시된 프로그램 모듈은, 또는 그 부분, 원격 메모리 스토리지 장치 내에 저장될 수 있다. 따라서, 예를 들어, 도 1에 도시된 바와 같이, 원격 애플리케이션 프로그램(189)은 원격 컴퓨터(182)의 메모리 장치에 상주할 수 있다. 도시된 네트워크 접속은 예시적이며, 컴퓨터들 간의 통신 링크를 확립하는 다른 수단이 사용될 수 있다.

개선된 레이트/디스토션 최적화 기술:

이하, 이미지 데이터 및/또는 비디오 데이터를 초기에 압축/인코딩 및 압축풀기/디코딩하도록 구성되는 어떤 예시적인 방법과 장치가 기재되지만, 데이터 압축 분야의 당업자는 제시된 기술들이 다른 유형의 데이터를 압축/인코딩 및 압축풀기/디코딩하기에 적합하고 이를 위해 채용될 수 있음을 인식할 것이다. 예를 들어, 어떤 방법과 장치는 오디오 데이터, 스피치 데이터 등을 압축/인코딩하는 데 사용하기에 적합할 수 있다.

또한, 예시적인 방법과 장치는 컴퓨터 내의 로직으로 구성될 수 있지만, 당업자는 이러한 방법 및 장치는 다른 유형의 장치, 기기 등으로 구현될 수 있음을 인식할 것이다. 여기서 사용되는 '로직'이라는 용어는 하드웨어, 펌웨어, 소프트웨어, 또는 이들의 임의의 결합, 및 원하는 기능을 전부 또는 부분 충족하도록 요구될 수 있는 것으로서 임의의 다른 지원 하드웨어 또는 다른 메커니즘을 포함하는 의미를 갖는다.

이를 감안하여, 데이터의 처리를 지원하는 어떤 형태의 로직에서 구현될 수 있는 여러 예시적인 방식이 제공된다.

본 발명의 일 태양에 따라, 비디오 및/또는 이미지 인코딩/디코딩 시스템의 성능을 향상시키기 위한 여러 신규한 기술이 제공된다. 어떤 예시적인 구현예에서, 이들 기술은 JPEG 및 JVT(조인트 비디오 팀) 표준(예를 들어, H.264/AVC) 등의 이미지/비디오 코딩 표준을 사용하여 채용된다. 예를 들어, JVT의 경우, 모션 정보 등의 특정 비디오 파라미터를 예측하기 위해 이용될 수 있는 서로 다른 예측 유형의 적용 선택이 동작가능하게 할 수 있는 구문 변경이 제공된다.

레이트 디스토션 최적화(RDO) 기술은, 종래의 인코딩 방법과 비교하여 인코딩 효율을 상당히 개선할 수 있기 때문에, 비디오와 이미지 인코딩/디코딩 시스템에서 매우 인기있다. 보다 높은 인코딩 효율을 제공하기 위한 지속적인 요구가 있다. 본 명세서는 기존 시스템에 비해 개선된 코딩 효율을 달성하기 위해 RDO 또는 다른 유사 기술의 성능을 상당히 향상시킬 수 있는 방법과 장치를 기술한다.

본 발명의 또다른 태양에 따라, 이미지 또는 매크로블록의 특성에 기초하여 어댑티브 및/또는 가변 크기 테드존 양자화를 사용함으로써, 예를 들어, 에지 및/또는 텍스처 검출 등의 이미지 선분석(pre-analysis)을 사용하여 RDO 또는 다른 유사 기술들을 더욱 결합 개선시킬 수 있다.

또한, 본 명세서는 프레임 레벨에서 서로 다른 예측 방식의 결합을 가능하도록 구현될 수 있는 몇몇 부가적인 예시적 구문 변화(syntax change)를 소개/정의함으로써, 비디오 코딩 방식의 성능을 더욱 향상시킨다. 예를 들어, 시퀀스의 몇몇 프레임 또는 부분에 있어서, 모션이 공간 도메인(spatial domain)보다 시간 도메인(temporal domain)에 더 상관될 수 있다 -또는 그 역(vice versa)-는 사실에 근거하여 하나의 기술이 발생한다. 이는, 프레임의 선분석을 이행함으로써, 또한, 가능한 한 적어도 2개의 방법을 이용하여 RDO 및/또는 RDO 유사 센스에서 보다 바람직한 방식을 선택하여 동일 프레임을 인코딩함으로써 이용될 수 있다. 따라서, 바람직한 방법은, 예를 들어, 디코더가 적절하게 프레임 을 디코드할 수 있도록, 예를 들어, 프레임 헤더 내에서의 결과적 데이터에 전달된다. 이 때, 예를 들어, 이러한 예시적

인 방법들 중 하나는, 공간 예측 또는 시간 예측을 이용하여 B프레임 내의 다이렉트 모드, 또는 공간 예측 모션 벡터 파라미터(spatial prediction motion vector parameter)나 시간 예측 모션 벡터 파라미터(temporal prediction motion vector parameter), 또는 심지어 0을 이용하여 P 프레임 내에 스킵 모드 모션 벡터 선택의 변경가능성을 포함할 수 있다.

하나 이상의 특성(파라미터)에 따라 어댑티브 및/또는 가변 데드존 양자화를 데이터에 적용:

이미지 데이터 압축 시스템에서, 예를 들어, 블록 또는 매크로블록과 같은 압축될 데이터는 다른 정보(데이터)와 비교하여 언제 서로 달리 코딩되어야 하는(예를 들어, 보다 높은 우선순위에서, 보다 높은 품질로, 등등) 보다 중요정보를 실제 포함할 수 있다. 이를 달성하는 하나의 방법은 서로 다른 양자화기 값을 이용하는 것이다. 예를 들어, 특정 구현 예에서, 보다 작은 양자화기 값은 '보다 중요 정보'에 대하여 이용될 수 있으며, 보다 큰 양자화기 값은 '보다 덜 중요한 정보'에 대하여 이용될 수 있다. 그러나, 그렇게 행하는 것은, 통상, 추후 압축풀기/디코딩이 성공적이도록 각각의 블록, 매크로블록, 매크로블록의 그룹 등에 대하여 이용되는 각각의 양자화기 값을 식별하는 정보의 전송을 요구한다. 불행히도, 이러 한 추가 정보는 압축 오버헤드(compressed overhead)와 인코더의 복잡도(complexity)를 증가시키는 경향이 있다. 따라서, 효율을 증가시키는 대신, 실제로는 효율이 감소될 수 있다.

도 2a 내지 도 2d를 참조하면, 어떤 예시적인 이미지/비디오 코딩 방식 내에 채용가능한 어떤 예시적인 데드존 양자화를 도시하는 그래프를 각각 나타낸다. 도 2a 내지 도 2d의 각각의 도면에서, 수직(y축)은 양자화된 값을 수평(x축)은 오리지널 값을 나타낸다. 도시된 예시적인 데드존 양자화기(도 2a의 202), 예를 들어, 균일 양자화기(uniform Quantizer)에 비해 코딩 효율을 상당히 향상시킬 수 있는 데드존 양자화 A에 관련된다. 종래의 데드존 양자화기는 종종 그 양자화 프로세스를 통하여 일정 또는 균일하게 유지되어, 데이터 내에 모든 기존 리던던시(redundancy)를 완전하게 이용할 수 있는 없었다.

데이터 내의 특정 정보의 중요성을 고려함으로써 그리고 블록/매크로블록 기반의 데드존 양자화기를 적용/변경함으로써, 예를 들어, 코딩 효율의 향상이 달성될 수 있다. 이는, 예를 들어, 각각의 양자화 빈(bin)(예를 들어, x축을 따라)의 커버리지를 적용함으로써, 그러나, 재구성 값(reconstructed value)을 변경시키지 않고, 행해질 수 있다. 예를 들어, 데드존 양자화기(202)를 도 2b 내지 도 2d의 데드존 양자화기(204, 206, 208)와 각각 비교해보자. 여기서, 재구성 값은, 물론 양자화 파라미터의 변경이 전달되지 않는다면, 코드된 데이터에 걸쳐 일정하게 유지될 것이다. 예를 들어, 제로 빈(도 2c)를 증가시킴으로써, 보다 많은 데이터가 이에 할당되어, 압축 방식에 따라 보다 높은 압축을 달성할 것이다. 재구성 값이 동일하게 유지되기 때문에 양자화기 내의 변화를 전달하기 위한 디코더에 대한 요구가 없음은 명백하다. 비록 이러한 것이 양자화 성능을 손상시킬 수 있다고 주장할 수 있지만, 그 처리가 특정 조건이 만족될 때 선택적으로 행해진다면, 예를 들어, 압축이 새로운 양자화기를 사용하여 초래된 디스토션보다 상당히 높게 되면, 반드시 그러한 것만은 아니다.

도 2b에서, 예를 들어, 데드존 양자화기는 제1 양/음 빈에 대한 0 빈만을 영향을 주는 반면, 다른 모든 빈은 동일하게 유지되고 사용된 재구성 값에 변화가 없다. 데드존의 어댑티브 추정을 사용하는 어댑티브 방식도 또한 가능하다(예를 들어, 레이트 디스토션 최적화 및 어댑티브 추정).

본 발명의 특정 예시적인 구현예에 따라, 양자화 선택 로직이 적어도 하나의 특성 또는 파라미터에 기초하여 여러 데드존 양자화기(또는 양자화 값)들 중에서 선택하기 위해 압축/인코딩 시스템 내에 제공된다. 예를 들어, 로직은 특정 이미지 특성에 따라 여러 데드존 양자화기(202, 204, 206, 및 208) 사이에서 선택할 수 있다.

특히, 그 로직은 이미지 데이터 내의 텍스처 및/또는 에지를 '보다 중요한' 정보를 나타내는 것으로 특성화하고, 이에 따라 높은 레벨의 품질을 제공하기 위한 방식으로 이러한 데이터를 코드하도록 구성될 수 있다. 로직은 이러한 판정 프로세스를 지원하기 위해 종래 텍스처 분석/탐지 및 에지 탐지 알고리즘을 사용할 수 있다.

이를 감안하여, 이하, 도 3을 참조하며, 이는 본 발명의 일 구현예에 따라 서로 다른 양자화 프로세스를 선택적으로 적용하기 위한 방법(300)을 나타내는 흐름도이다. 동작(302)에서, 초기 데이터 세트는 복수개의 데이터 부분으로 분할된다. 예를 들어, 이미지 또는 비디오 데이터는 복수개의 블록, 매크로블록, 슬릿, 슬라이스, 섹션 등으로 분할될 수 있다. 동작(304)에서, 동작(302)으로부터의 복수개의 데이터 부분들 중 적어도 하나는, 예를 들어, 적어도 하나의 특성 또는 파라미터에 기초하는 어떤 방식으로 분류된다. 예를 들어, 그 분류는 에지 특성, 텍스처 특성, 평활 특성, 루미넌스 특성, 크로미넌스 특성, 컬러 특성, 노이즈 특성, 오브젝트 특성, 모션 특성, 사용자 선호도 특성, 사용자 인터페이스 포커스 특성, 레이아웃 특성, 타이밍 특성, 볼륨 특성, 주파수 특성, 피치 특성, 톤 특성, 품질 특성, 비트 레이트 특성, 데이터형 특성, 해상도 특성, 암호화 특성 등에 기초할 수 있다.

동작(306)에서, 동작(304)에서 분류된 부분은 대응 양자화 데이터를 산출하기 위해서 복수개의 양자화 프로세스 중 적어도 2개를 이용하여 처리된다. 동작(308)에서, 동작(306)으로부터의 양자화 데이터들 중의 하나가, 예를 들어, 적어도 적어도 하나의 임계치를 충족시키거나 동작(304)에 사용된 분류 파라미터에 관련되는 측정에 기초하여 선택되

는 판정이 행해진다. 동작(310)에서, 동작(308)으로부터의 양자화 데이터는 추후 디코드될 수 있는 방식으로 인코딩된다.

방법(300)을 사용하여, 이미지 또는 비디오 부분들을 근본적으로 분석하여 특정 부분들을 다른 것보다 보다 중요한 것으로 여기는 로직이 제공될 수 있다. 따라서, 여러 데드존 양자화기가 보다 중요한 데이터 부분에 적용되어, 그 결과, 데드존 양자화기(들)가 이렇게 중요한 데이터에 대하여 원하는 임계치 요건을 충족하는 지를 판단하도록 양자화 데이터가 분석된다. 예를 들어, 품질 또는 노이즈 임계치 요건이 실행될 수 있다.

예를 들어, 비디오 시퀀스에서, 예지는 종종 모션 보상 기술에서 상당히 중요한 역할을 한다. 이는, 예를 들어, 이하 보다 상세히 설명하는 도 7에 나타난 블록도에서 설명된다. 기본적으로, 복수개의 데드존 양자화기 중에 어느 것이 가장 사용하기에 적합한지를 판단하는 분석에 따라, 예를 들어, 레이트 디스토션 최적화(RDO) 기준에 따라, 이미지(프레임) 또는 이미지의 부분에 대한 이미지 분석을 이행하는 것 가능하다.

도 4를 참조하면, 로직(400)은 본 발명의 일 구현예에 따라 온 더 플라이 방식(on-the-fly) 판정 프로세스를 나타낸다. 여기서, 블록(402)에서 입력 프레임/이미지(또는 그 부분)은 이미지 분석된다. 이 예에서, 이미지 분석은 에지 탐지 및/또는 텍스처 분석 능력을 포함하며, 그 출력은 데드존 판정 블록(404)에 제공된다. 그 후, 데드존 판정 블록(404)은 인터도(406)가 입력 프레임/이미지(또는 그 부분)을 인코딩할 때 특정 또는 달리 선택된 데드존 양자화기를 사용하도록 할 수 있다.

데드존 양자화에 대한 온 더 플라이 방식 판정을 행하는 대신, 도 5 내의 로직(500)은, 복수개의 데드존 양자화기가 사용되어 특정 임계치 기준에 기초하여 RDO 판정이 행해지는 방법(300)을 지원하도록 구성될 수 있다. 여기서, 입력 매크로블록(MB)/이미지(502; 또는 유사 부분)는 여러 데드존 인코딩 블록(504, 506, 508 및/또는 510)에 (선택적으로) 제공되며, 이들 여러 데드존 인코딩 블록으로부터의 출력은 RDO 판정블록(512)과 선택된 인코딩 데이터 출력에서 분석된다. 이 예에서 설명한 바와 같이, 데드존 인코딩 블록/프로세스의 부분 또는 전부는 병렬로 발생할 수 있다. 다른 구현예에서, 이러한 처리는 직렬로 발생하도록 타이밍될 수 있다.

도 6a 내지 도 6b는 선택적 기록 판정처리가 이용되는 로직을 도시한다. 도 6a에서, 예를 들어, 입력 MB/이미지(502)가 데드존 인코딩 블록(602, 604)에 제공된다. 기록 판정 블록(606)은 데드존 인코딩 블록(602)으로부터의 출력을 고려하며, 데드존 인코딩 블록(602, 604)의 출력들 간의 선택(608)에 영향을 준다. 기록 판정 블록(606)은 또한 데드존 인코딩 블록(604)을 선택적으로 초기화할 수 있다. 도 6b에서, 로직(610)은 로직(600)과 유사하지만 기록 판정 블록(606)에 관련된 선택(608)을 갖는 대신, 데드존 인코딩 블록(602, 604)으로부터의 출력을 분석하여 어느 것을 출력할 지를 판정하도록 구성된 RDO 판정 블록(612)을 포함한다.

도 6a와 도 6b의 기록 판정 블록(606)은 여러 기준에 기초하여 판정을 하도록 구성될 수 있다. 예를 들어, 특정 구현예에서, 품질 제한, 레이트 제한 및/또는 다른 RDO 개념/임계치가 고려될 수 있다.

예를 들어, 레이트/품질이 특정 임계치 이상인 것과 같이, 종전에 정의된 어떤 조건이 충전하기만 하면, 추가적인 데드존 양자화기를 사용가능 것이 가능하다. 특정 예시적인 실험에서, 오리지널 값보다 약 30% 이상 큰 데드존을 구비한 데드존 양자화기가 성공적으로 선택되었다. 다른 것으로, AC 주파수 또는 에지 유형과 같은 이미지 특성에 따라 어댑티브가능한, 데드존 양자화기가 또한 사용될 수 있다.

도 7에서, 예시적인 인코딩 로직(700)은 어떻게 초기 이미지(702; 또는 그 부분)가 이미지 분석 처리(704)에 따라 처리되어, 이 예에서는, 에지 탐지 데이터(706)를 생성하는지를 설명한다. 그 후, 에지 탐지 데이터(706)은 블록 분류 처리(708)에 제공되어 블록 분류 데이터(710)을 생성한다. 그 후, 블록 분류 데이터(710)는 초기 이미지(702)를 따라 이미지 인코딩 프로세스(712)에 제공되며, 이는 그 후 인코딩된 이미지(714)를 생성한다. 이는 이미지에 대한 단순 인코딩 처리의 일 예이다. 여기서, 이미지는 분석(예를 들어, 에지 탐지 알고리즘을 사용하여)된 후, 블록이 이러한 정보에 따라 분류된다. 근본적으로, 매크로블록의 위치(i,j)에서 할당된 양자화기(QP)에 대한 적절한 데드존(DZ)을 추후에 할당하는 N-ary 매트릭스(DZ_p)가 정의된다(N은 정의된 데드존 개수에 의존).

논 레지듀얼 모드(예를 들어, B 프레임 다이렉트 모드에 대하여)의 고려

비록 상기 설명이 정적 이미지에 주로 초점을 맞추어 기재하였으나, 동일한 개념이 또한 비디오에도 용이하게 적용될 수 있다. 예를 들어, 이미지 선분석 및 데드존 양자화기는 비디오 압축에서도 다소 직접적으로(straightforward) 사용될 수 있다.

대단히 흥미롭고 상기 개념으로부터 상당한 이점을 얻을 수 있는 일례로 B프레임 내의 다이렉트 모드의 사용을 들 수 있다. 다이렉트 모드는 기본적으로 모션 파라미터의 어떠한 전송도 요구하지 않는 특정 모드인 데, 이는 이러한 것이 시간 또는 공간 예측을 통해 집적 예측될 수 있기 때문이다. 다이렉트 모드에 대한 추가적인 정보가 함께 계류중인 미

국 가출원번호 제 60/385,965호에서 발견될 수 있다.

만약 전송될 레지듀(residue)가 없으면, 본 발명의 다른 양태에 따라, 여기서 는 레지듀얼 다이렉트 모드(Non Residual Direct Mode)로서 지칭되는 특정 모드의 사용에 의해 효율이 더욱 향상될 수 있다. 후술하는 바대로, 는 레지듀얼 다이렉트 모드는 런 렱스 인코딩(RLC; Run Length enCoding) 방식을 이용하도록 구성될 수 있다. 여기서, 예를 들어, 초래되는 디스토션이 충분히 작고 보다 높은 효율의 RLC에 기인한 비트 레이트 감소가 중요해지면, 데드존 양자화기가 원하는 솔루션을 제공할 수 있다. 기본적인 아이디어는, 특정 조건하에서, 비록 레지듀가 존재하더라도 그 레지듀 없이 코드되도록 다이렉트 모드(Direct Mode)를 선택적으로 하거나 코드되도록하는 것이다. 도 5와 동일한 개념에 기초한 방식은, 예를 들어, 도9에 나타난 것과 같이, 구현될 수 있으며, 또한 모든 RDO 처리 내에서 는 레지듀얼 다이렉트 모드가 다른 가용 모드와 비교되어 검사된다. 특정 사례에서, 이러한 방식의 성능은, 사용된 RDO가 이러한 경우에 대하여 부적절하기 때문에, 예측된 것 만큼 양호하지 않을 수 있다. 양자화 값에 의존하는 다른 처리는, 예를 들어, 블로킹 구조물(blocking artifact)을 제거하는 데 사용되는 인루프 필터(in-loop filter; 도시 생략) 등에 영향을 받는다. 특히, 비록 성능이 더 낮은 비트 레이트에서는 양호해 보일지라도, 보다 높은 비트 레이트에서는 상당히 열화하면, 는 레지듀얼 다이렉트 모드의 고려없이 보다 큰 양자화기의 사용에 의해 성능이 향상될 수 있다.

이미지의 예에서 행해진 것과 유사하게, 본 발명의 특정 양태에 따라서, 종전에 정의된 어떤 조건들이 충족되기만 하면, 그리고 특히, 다이렉트 모드에 관련된 레지듀가 중요하지 않으면, 이러한 모드를 고려하도록 로직이 구성될 수 있다. 그 로직은, 예를 들어, 다이렉트 모드의 코드화된 블록 패턴(CBP; Coded Block Pattern)을 검사함으로써 이 레지듀가 얼마나 중요한 것인지를 추정 조건으로서 사용되도록 구성될 수 있다. CBP가, 크로미넌스 정보의 고려없이, 특정 임계치 이하이면, 이는 레지듀가 중요하지 않음을 암시하는 것으로 간주될 수 있고, 스킵되었으면, 과도한 디스토션을 초래하지 않을 것이다. 더욱이, 에지 및 텍스처의 부존재와 같이 다른 이미지 특성이 이러한 비디오 인코딩 처리/로직 내에 사용될 수 있다.

특정 구현예에서, 로직은 모든 루미넌스 레지듀가 손상되지 않으면서(intact) 크로미넌스 레지듀(crominance residue)만이 제거될 수 있는 지를 검사함으로써, 이를 보다 연장할 수 있다. 이러한 모드에 레지듀 전체 또는 크로미넌스를 전송하지 않는 것이, 보다 우수한 성능을 부여하는 지를 검사하는 것은, 이러한 개념을 매크로블록에 대한 모든 가능한 모드로 연장시킬 수 있다. 그러나, 이는 인코딩 복잡도를 훨씬 복잡하게 하는 것은 명확하다.

도 9를 참조하면, 로직(900)은, 입력 MB/이미지(902)가 서로 다른 다이렉트 모드(B 프레임) 또는 카피 모드(인터 프레임) 인코딩 처리 블록(902, 904, 906, 및 908)에 어떻게 제공되는지와 이들 처리블록으로부터의 출력이 적절한 출력을 선택 하는 RDO 판정 블록(910)에 기초하여 어떻게 프레임에 제공되는지를 설명한다.

도 10을 참조하면, 또한, 로직(1000)이 어떻게 방식 판정 블록(1002)과 선택(1004)이 사용자 입력, 애플리케이션, 시스템 요건 등에 의존하여 추가적인 선택가능성을 제공하도록 포함되는 지를 더 설명한다. 여기서, 방식 판정 블록(1002)은 하나 이상의 다이렉트 모드(B 프레임) 또는 카피 모드(인터 프레임) 인코딩 처리 블록(902, 904, 906, 및 908)에 선택적으로 제공한다. 선택(1004)은 방식 판정 블록(1002) 또는 다른 처리/로직/입력에 의해 제어될 수 있다.

B 프레임에 대한 라그랑지 곱수의 예시적인 사용

RDO 기술은 종종 라그랑지 곱수(Lagrangian multiplier)의 개념에 기초한다. 예를 들어, 디스토션과 비트 레이트를 함께 최소화하는 특정 모드가 선택될 수 있다.

이러한 함수는 최소화로서

$$J(\text{Mode}|QP, \lambda) = SSD(\text{Mode}|QP) + \lambda \cdot R(\text{Mode}|QP)$$

으로 표현되며, 여기서, QP는 매크로블록 양자화기, λ 는 모드 판정을 위한 라그랑지 곱수이며, *Mode* 는 RDO 처리 내에 검사되어 가능하면 선택되는 매크로블록 모드를 나타낸다.

예를 들어, 특정 구현예에서, 라그랑지 곱수 λ 는 인터 프레임 또는 인트라 프레임에 대하여

$$\lambda_{I,P} = 0.85 \times 2^{\frac{QP}{3}} \quad \text{또는}$$

$$\lambda_{I,P} = 5 \times \frac{QP+5}{34-QP} \times \exp^{\frac{QP}{10}}$$

으로 선택될 수 있으며, 그 반면에, B프레임에서는, JVT와 같은 대부분의 코덱에서, 이는 $\lambda_B = 4 \times \lambda_{I,P}$ 으로서 선택된다.

λ 의 추가 가중치는, 보다 낮은 오버헤드 모드, 예컨대 다이렉트 모드는 RDO에 대하여 여전히 매우 우수한 성능을 제공할 수 있는 반면에, 특히 B프레임에 있어서는, 모드들이 전송된 다수의 모션 정보에 기인하여 대용량의 오버헤드를 가질 수 있기 때문에, 그 오버헤드 모드를 낮추기 위한 선호도를 부여하기 위해 행해진다.

그러나, 본 발명의 따른 특정 실험에 기초하여, 그 가중치는 상술한 바와 같이 상수이면 안되지만, 그 대신, 양자화기 QP값에 다시 의존하여야 함이 발견되었다.

특히, $\lambda_B = f(QP) \times \lambda_{LP}$ 이라고 정의하면, 이들 실험으로부터, $f(QP)=4$ 로 고정된 경우보다 훨씬 우수한 압축효율로 이용되는 2개의 $f(QP)$ 함수는

$$f(QP) = \max(2, \min(4, \frac{QP}{6})) \text{ 와}$$

$$f(QP) = \max(2, \min(4, \frac{QP+2}{8})) \text{ 이다.}$$

또한, 이러한 관측은 매우 높은 λ 을 가짐으로써, 예를 들어, 모션 예측의 정확성과 같은 다른 중요 정보의 정확성에 영향을 주게 되어, 그 결과 주변 매크로 블록의 인코딩에 부정적인 영향을 미친다는 사실에 기인한다.

예측 의존 MB 모드의 어댑티브 유형 선택의 사용

함께 계류중인 미국 가출원번호 제60/385,965호, 제60/376,005호, 미국 출원번호 제10/186284호에 기재된 바와 같이, 시퀀스와 프레임은 서로 다른 유형의 지배적 모션 상관관계를 가질 수 있다.

특히, 정적 배경에서 일정 속성을 갖는 소형 객체에 있어서, 시간적으로 인접한 프레임(시간 도메인)으로부터 모션 벡터(MVs)의 사용은 보다 우수한 예측을 이행하고, 보다 높은 성능을 낼 수 있게 한다. 부드러운 모션의 대형 객체는 공간 도메인(인접 매크로블록에서 보다 높은 상관관계를 갖는 반면, 다른 경우에는, 시간 도메인과 공간 도메인으로부터의 정보가 모두 예측을 위해 중요할 수 있다. 이들 유형의 상관관계는, 예를 들어, B와 P 프레임에서 다이렉트 모드와 P프레임에서의 모션 벡터 예측기에 대한 스킵 내에서 부분적으로 이용될 수 있다. P 프레임에서의 모션 벡터 예측기에 대한 스킵(Skip on Motion Vector Predictor)에 대하여 보다 많은 정보는, 예를 들어, Jani Lainema 및 Mart a Karczewicz가 2002년 5월, JVT Meeting, Fairfax에서 '스킵 모드 모션 보상' 문서 JVT-C027을 참조하면 된다. 따라서, 로직이 어느 유형의 예측이 서로 다른 프레임에서 우세한 지를 어느 정도 전달하면, 매우 높은 성능이 달성될 수 있다.

따라서, 본 발명의 특정 양태에 따라, 인코딩 로직은, 예측 의존 모드에 대한 예측 방식이 사용되는 프레임, 슬라이스 또는 다른 유사 레벨에서 어떤 방식으로 전달 또는 식별하도록 구성된다. 이를 달성하기 위한 하나의 예시적 구문은, 예를 들어 JVT 내에서, 도 8의 차트 800에 제공된다. 이 구문은, 물론, 다른 인코딩 디자인에서 변경/상이하하다.

차트 800에서의 예시적인 구문에 있어서, 가능한 3개의 예측의 경우가 P 프레임 스킵 모드에 대하여 허용된다. 이 예에서, 모션 카피 예측 모드(Motion-Copy prediction mode), 시간 예측 스킵 모드(temporal prediction skip mode), 및 제로 스킵 모드(zero skip mode)이다. 이들 경우의 각각은 {0.1.2} 사이의 값으로 할당되며, 이는 고정 크기 코드워드, 이 예에서 $u(2)=2$ 를 사용하여 코드될 수 있거나, 심지어 엔트로피 코딩(entropy coding: UVLC 또는 CABAC을 이용하는 $e(v)$)된다. 당업자는 이런 채용될 수 있는 정보를 전달하는 다른 방식이 있음을 인식할 것이다. 차트 800에서, B 프레임에 대하여 어댑티브 공간/시공간 고려(Adaptive Spatial/Spatio-temporal Consideration)(direct_mv_spatial)와 인터 프레임 내에 어댑티브 카피/모션 카피 스킵 모드(adaptvie Copy/Motion-Copy Skip Mode)(copy_mv_spatial)에 더하여 JVT 내에 픽처 레이어 RBSP 구문(picture layer RBSP syntax)을 설명한다. 단지 이들 두 모두가 사용되면, 정의된 디스크립터는 단지 1 비트만을 취하여, $u(n=1)$ 이지만, 보다 많은 케이스가 사용되면(제로 바이어스(zero bias)의 공간 예측 또는 정적 시간 예측의 고려), 보다 많은 비트가 할당되며($n>1$) 심지어 $e(v)$ 에 의해 정의된 파라미터에 대하여 엔트로피 코딩을 사용할 수 있다.

B 프레임에 있어서, 이 예에서는, 공간 모션 벡터 예측은 모든 다이렉트 모드 모션 파라미터에 대하여 하나의 모드로서, 시간적으로 예측된 파라미터는 제2 모드로서 사용될 수 있다. 기재된 다른 다이렉트 모드, 예를 들어, 함께 계류중인 미국 가출원번호 제60/385,965호가 또한 고려/포함될 수 있다. 인코딩 로직은 어느 예측 모드가 프레임 또는 슬라이스 레벨에서 이용될 수 있는 지를 전달하도록 구성된다. 그 부분은, 예를 들어, RDO 기반 방식(예를 들어, 도 9)을 사용하여 이행될 수 있다. 또한, 특정 구현예에서, 인코딩 로직은 인코더 및/또는 디코더 상의 특정 요건에 기인하여 명시적으로 특정 모드를 사용할 수도 있다. 특정 예에서, 통상, 공간 예측이 간단하게 계산 상 보다 간단한 것을 고려하면(예를 들어, 분할, 모션 벡터의 저장장치를 요구하지 않고, 타이밍 정보에 독립적이다), 공간 예측은 부분 애플리케이션(예를 들어, 도 10)에 대해서 바람직한 선택일 수 있다.

다른 구현예에서, 이러한 문제점들이 쟁점이 아니라면, 그 조합은 더욱 향상된 인코딩 성능을 산출할 수 있다. 인코딩된 시퀀스의 일례는 도 11의 예시적인 도면에 나타난다. 여기서, P 및 B 프레임은 화면 변경(scene change)에 따라 도시된다. 화살표에 의해 나타난 바와 같이, 서로 달리 전달된 P 및 B 프레임은 대응 스킵, 또는 다이렉트 모드 매크로블록에 대하여 도시된다. 또한, 전달링은 인코딩 로직이 모션 벡터 코딩에 대하여, 또는 다른 모드(예를 들어, 함께 계류 중인 미국 가출원 번호 제60/376,005호, 미국출원번호 제10/186,284호)의 예측에 대하여 어떻게 모션 벡터 예측을 이행하는지의 표시일 수 있다.

도 11에 도시된 바와 같이, 서로 다른 프레임들은 그들의 대응 다이렉트(B) 및 스킵(P) 모드에 대하여 서로 다른 유형의 예측을 전달한다. P_z , P_T , 및 P_M 은 예를 들어, 영의, 시간 및 공간 예측(모션 카피) 예측에 대하여 정의하며, 및 B_T 및 B_{sp} 는 다이렉트 모드에 대한 시간 및 공간 예측을 정의한다.

정적 시간/공간 블록을 고려하는 타임 스탬프 독립 다이렉트 모드

여러 유형의 예측은, 특히, B 프레임에서 다이렉트 모드에 있어서, 서로 다른 유형의 모션 및 시퀀스에 대하여 보다 적합할 수 있다. 시간 또는 공간 예측만을 이용하여, 몇몇의 경우는 허용가능한 성능을 제공할 수 있지만, 다른 성능은 상당히 열악할 수 있다. 상술한 바와 같이, 또는 함께 계류 중인 미국 가출원번호 제60/385,965가 제공되는 경우에서, 솔루션은 훨씬 더 우수한 성능을 제공할 수 있다.

예를 들어, 매우 효율적이며 시간 및 공간 예측기 양자의 성능을 결합하는 또다른 경우가 제공되며, 분할 및/또는 독립적인 타이밍 요구하지 않음으로써 공간 예측을 간단하게 유지하는 경향이 있다.

특정 구현예에서, 공간 예측은 시간예측보다 유용하다(예를 들어, 속성에 기인하여). 따라서, 예를 들어, 공간 예측은 다이렉트 모드의 주 예측으로서 이용된다. 한가지 가능한 예외는 시간 예측기로부터의 모션 정보와 참조 프레임이 0인 경우이다. 이러한 경우, 다이렉트 모드의 대응 블록에 대한 모션 정보와 참조 프레임도 0으로 간주된다. 더욱이, 공간 예측은 공간 0 바이어스됨(spatial zero biased-ness) 및/또는 정적 서브파티션(stationary subpartition)을 또한 고려하여 세련(refine)되어질 수 있다. 따라서, 현재 예측된 블록에 대한 인접 매크로블록 또는 블록의 전부 또는 부분이 0 모션(또는 매우 가까운(예를 들어, 정수 모션 벡터가 0))과 참조 프레임을 가지면, 또한 전부 또는 부분의 매크로블록이 0의 모션을 갖는 것으로 간주된다. 이들 개념은 모두 정적 배경을 보호하는데, 특히, 이동 객체의 에지에서 이러한 조건이 되지 않으면 매우 디스토션될 수 있는 것으로부터 보호할 수 있게 한다.

16×16 매크로블록의 경우에 대하여 간략화된 흐름도(1200)가 도 12에 도시된다. 다이렉트 모드에 대한 공간 예측을 사용하여 이러한 개념은 훨씬 작은 블록(예를 들어, 8×8 또는 4×4)(또는 대형 블록 또는 다른 형태의 부분)에 또한 연장될 수 있다. 동작(1202)에서, 공간 예측기(MV_a , MV_b , MV_c) 및 시간 예측기(MV_t)가 동작(1204)에 제공되며, MV_{Direct} 는 중간값(MV_a , MV_b , MV_c)으로 설정된다. 동작(1206)에서, MV_{Direct} 가 0으로 설정되는 동작(1208) 또는 추가 판정을 위한 동작(1210)으로 진행되는 MV_t 에 기초하여 판정이 행해진다. 동작(1210)의 결과, MV_{Direct} 는 동작(1208)에서 0으로 설정되거나 변경되지 않으며, MV_{Direct} 는 출력이다.

본 발명의 다양한 방법 및 장치의 어떤 바람직한 구현예가 첨부한 도면에 도시되고 상기 상세한 설명에서 기재되었지만, 본 발명은 개시된 예시적인 실시형태에 한정하는 것이 아니고, 본 발명의 정신으로부터 벗어남이 없이 다수의 재배치, 변형 및 대체가 가능하다는 것이 인식될 것이다.

발명의 효과

본 명세서에서, 이미지, 비디오 시퀀스, 또는 다른 유형의 데이터를 인코딩하기 위한 레이트 디스토션 최적화에 관련된 여러 개념들이 제공되었다. 여러 다른 예측의 경우, 특히, 이미지의 초기 개시에서 전달될 수 있는 P 및 B 프레임 내에서 각각 스킵 및 다이렉트 모드 등의 경우가 허용됨으로써, 상당히 높은 압축 효율을 가능하게 하는 비디오 시퀀스 내에 채택될 수 있는 추가적인 구분이 설명되었다. 정적 시간 예측기 뿐만 아니라 공간 모션 벡터 예측을 고려하는, B 프레임에 대한 고효율의 타임 스탬프 독립 다이렉트 모드(time stamp independent Direct Mode)가 제공되었다. 상기 방법 및 장치의 전부 또는 부분은 여러 이미지/비디오/데이터 코딩 시스템의 성능을 상당히 향상시키도록 구현될 수 있다.

(57) 청구의 범위

청구항 1.

복수개의 서로 다른 양자화 함수(quantization function)를 사용하여 데이터의 적어도 일부분을 처리하여 복수개의 대응하여 양자화된(corresponding quantized) 데이터 부분을 생성하는 단계; 및

적어도 하나의 임계치에 기초하여 상기 양자화된 데이터 부분들 중의 하나를 선택적으로 출력하는 단계를 포함하는 방법.

청구항 2.

제1항에 있어서,

초기 데이터를 상기 데이터의 적어도 일부분을 포함하여 복수개의 부분으로 분할하는 단계; 및

상기 데이터의 적어도 일부분을 적어도 하나의 분류 특성(classification characteristic)에 기초하여 분류하는 단계를 더 포함하는 방법.

청구항 3.

제2항에 있어서,

상기 적어도 하나의 임계치는 상기 적어도 하나의 분류 특성에 관련되는 방법.

청구항 4.

제2항에 있어서,

상기 초기 데이터는 이미지 데이터, 비디오 데이터, 오디오 데이터, 및 스피치 데이터(speech data)를 포함하는 그룹으로부터 선택되는 데이터를 포함하는 방법.

청구항 5.

제2항에 있어서,

상기 초기 데이터는 이미지 데이터 또는 비디오 데이터를 포함하며, 상기 적어도 일부분은 블록, 매크로블록(macroblock), 슬릿(slit), 슬라이스(slice), 및 섹션(section)을 포함하는 그룹으로부터 선택되는 방법.

청구항 6.

제2항에 있어서,

상기 적어도 하나의 분류 특성은, 에지(edge) 특성, 텍스처(texture) 특성, 평활(smoothness) 특성, 루미넌스(luminance) 특성, 크로미넌스(crominance) 특성, 컬러 특성, 노이즈 특성, 오브젝트 특성, 모션(motion) 특성, 사용자 선호도 특성, 사용자 인터페이스 포커스(user interface focus) 특성, 레이어링(layering) 특성, 타이밍 특성, 볼륨(volume) 특성, 주파수 특성, 피치(pitch) 특성, 톤(tone) 특성, 품질 특성, 비트 레이트(bit rate) 특성, 데이터형 특성, 해상도(resolution) 특성, 및 암호화 특성을 포함하는 특성들의 그룹으로부터 선택되는 방법.

청구항 7.

제1항에 있어서,

상기 복수개의 서로 다른 양자화 함수는 적어도 2개의 서로 달리 동작하는 데드존 양자화기(Deadzone Quantizer)를 포함하는 방법.

청구항 8.

제7항에 있어서,

상기 데드존 양자화기들 중 적어도 하나는 어댑티브 커버리지 데드존 양자화기(adaptive coverage Deadzone Quantizer)를 포함하는 방법.

청구항 9.

제7항에 있어서,

상기 데드존 양자화기들 중 적어도 하나는 가변 커버리지 크기 데드존 양자화기(variabel coverage size Deadzone Quantizer)를 포함하는 방법.

청구항 10.

제1항에 있어서,

상기 양자화된 데이터 부분들 중 하나를 선택적으로 출력하는 단계는 상기 양자화된 부분들 중 상기 하나를 인코딩하는 단계를 더 포함하는 방법.

청구항 11.

제1항에 있어서,

상기 적어도 하나의 임계치에 기초하여 상기 양자화된 데이터 부분들 중 하나를 선택적으로 출력하는 단계는, 레이트 디스토션 최적화(RDO;rate distortion optimization)를 이행하여 상기 양자화된 데이터 부분들 중의 하나를 선택하는 단계를 더 포함하는 방법.

청구항 12.

이미지 데이터의 적어도 일부분에 대하여 적어도 하나의 특성 분석을 이행하는 단계;

상기 적어도 하나의 특성 분석에 기초하여 인코더 내에 적어도 하나의 어댑티브 양자화 파라미터(adaptive quantization parameter)를 선택적으로 설정하는 단계; 및

상기 인코더로 상기 이미지 데이터의 적어도 일부분을 인코딩하는 단계를 포함하는 방법.

청구항 13.

제12항에 있어서,

상기 적어도 하나의 특성분석은, 예지 특성, 텍스처 특성, 평활 특성, 루미넌스 특성, 크로미넌스 특성, 컬러 특성, 노이즈 특성, 오브젝트 특성, 모션 특성, 사용자 선호도 특성, 사용자 인터페이스 포커스 특성, 레이어링 특성, 타이밍 특성, 품질 특성, 비트 레이트 특성, 데이터형 특성, 및 해상도 특성을 포함하는 특성들의 그룹으로부터 선택되는 적어도 하나의 이미지 분석 특성을 고려하는 방법.

청구항 14.

제12항에 있어서,

상기 적어도 하나의 어댑티브 양자화 파라미터는 상기 인코더 내에 어댑티브 커버리지 데드존 양자화기와 관련되는 방법.

청구항 15.

제12항에 있어서,

상기 적어도 하나의 특성 분석에 기초하여 상기 인코더 내에 상기 적어도 하나의 어댑티브 양자화 파라미터를 선택적으로 설정하는 단계는, 비디오 프레임을 포함하는 상기 복수개의 이미지 데이터 부분들 각각에 대하여 온 더 플라이 방식으로(on-the-fly) 이행되는 방법.

청구항 16.

적어도 2개의 서로 다른 데드존 양자화기를 이용하여 이미지 데이터의 적어도 일부분을 인코딩하는 단계; 및

적어도 하나의 판정 팩터(decision factor)에 관련되는 레이트 디스토션 최적화(RDO)에 기초하여 상기 적어도 2개의 서로 다른 데드존 양자화기들 중 하나의 출력에서 바람직한 인코딩 데이터를 식별하는 단계를 포함하는 방법.

청구항 17.

제1 데드존 양자화기를 이용하여 이미지 데이터의 적어도 일부분을 인코딩하는 단계;

상기 제1 데드존 양자화기의 출력이 적어도 하나의 판정 팩터를 만족하는 지를 판정하는 단계; 및

만족하는 경우, 상기 제1 데드존 양자화기의 출력을 출력하며,

만족하지 않는 경우, 상기 제1 데드존 양자화기와 다른 적어도 제2 데드존 양자화기를 이용하여 상기 제1 이미지 데이터의 일부분을 인코딩하는 단계를 포함하는 방법.

청구항 18.

제17항에 있어서,

레이트 디스토션 최적화(RDO; rate distortion optimization) 판정에 기초하여 상기 이미지 데이터의 적어도 일부분의 허용가능한 인코딩 버전(acceptable encoded version)을 식별하는 단계를 더 포함하는 방법.

청구항 19.

이미지 데이터의 적어도 일부분에 대하여 이미지 분석을 이행하는 단계;

상기 분석된 이미지 데이터 부분에 대하여 블록 분류를 이행하는 단계;

상기 블록 분류된 이미지 데이터 부분의 데드존 양자화를 이행하는 단계; 및

상기 데드존 양자화된 이미지 데이터 부분의 인코딩을 이행하는 단계를 포함하는 방법.

청구항 20.

제19항에 있어서,

상기 이미지 분석은 에지 탐지 분석 및 텍스처 분석을 포함하는 그룹으로부터 선택되는 적어도 한 유형의 분석을 포함하는 방법.

청구항 21.

제19항에 있어서,

상기 블록 분류는 상기 이미지 분석에 기초하여 동작 구성되며, 상기 데드존 양자화는 상기 블록 분류에 기초하여 동작 구성되는 방법.

청구항 22.

적어도 2개의 서로 다른 인코더를 이용하여 비디오 이미지 데이터의 적어도 일부분을 인코딩하는 단계 -상기 2개의 서로 다른 인코더들 중 적어도 하나는 상기 비디오 이미지 데이터의 논 레지듀얼 모드(Non Residual mode)를 지원하도록 동작된 데드존 양자화기를 포함함-; 및

적어도 하나의 판정 팩터에 관련되는 레이트 디스토션 최적화(RDO) 판정에 기초하여 상기 적어도 2개의 서로 다른 인코더 중 하나의 출력에서 바람직한 인코드 프레임 데이터를 식별하는 단계를 포함하는 방법.

청구항 23.

제22항에 있어서,

상기 데드존 양자화기는 어댑티브하며, 상기 논 레지듀얼 모드를 선택적으로 사용하는 방법.

청구항 24.

제22항에 있어서,

상기 인코더들 중 적어도 하나는 상기 비디오 이미지 데이터의 다이렉트 모드(Direct Mode)를 지원하도록 동작 구성되는 방법.

청구항 25.

제22항에 있어서,

상기 인코더들 중 적어도 하나는 상기 비디오 이미지 데이터의 카피 모드(Copy Mode)를 지원하도록 동작 구성되는 방법.

청구항 26.

제22항에 있어서,

상기 논 레지듀얼 모드는 런 렱쓰 인코딩(RLC;Run Length enCoding) 방식에 기초하여 동작 구성되는 방법.

청구항 27.

제22항에 있어서,

상기 데드존 양자화기는 어댑티브하게 동작하는 방법.

청구항 28.

제22항에 있어서,

상기 데드존 양자화기는 상기 비디오 이미지 데이터의 적어도 일부분에 관련되는 적어도 하나의 특성에 기초하여 동작 구성되는 방법.

청구항 29.

제22항에 있어서,

상기 바람직한 인코드 프레임을 식별하는 단계는 상기 다이렉트 모드 프레임 데이터에 관련되는 레지듀(residue)의 양을 분석하는 단계를 더 포함하는 방법.

청구항 30.

제29항에 있어서,

상기 레지듀의 양을 분석하는 단계는 상기 다이렉트 모드 프레임 데이터의 코드화된 블록 패턴(CBP:Coded Block Pattern)을 검사하는 단계를 더 포함하는 방법.

청구항 31.

제22항에 있어서,

상기 데드존 양자화기를 구비한 상기 인코더는, 크로미넌스 레지듀 데이터(chrominance residue data), 루미넌스 레지듀 데이터(luminance residue data), 및 모든 레지듀 데이터를 포함하는 레지듀 데이터 그룹으로부터 선택되는 적어도 한 유형의 레지듀를 제거하도록 동작 구성되는 방법.

청구항 32.

이미지 데이터의 적어도 일부분의 적어도 하나의 특성에 기초하여 양자화 함수를 갖는 인코딩 로직을 동작 구성하는 적어도 하나의 라그랑지 곱수(Lagrangian multiplier)를 선택적으로 변경하는 단계; 및

상기 인코딩 로직을 이용하여 상기 이미지 데이터의 적어도 일부분을 인코딩하는 단계를 포함하는 방법.

청구항 33.

제32항에 있어서,

상기 양자화 함수는

$$J(\text{Mode}|\underline{QP}, \lambda) = \text{SSD}(\text{Mode}|\underline{QP}) + \lambda \cdot R(\text{Mode}|\underline{QP})$$

에 기초하는 매크로블록 양자화기를 포함하며, 여기서, λ 는 모드 판정에 대한 상기 라그랑지 곱수가며, Mode는 레이트 디스토션 최적화(RDO) 프로세스를 이용하여 검사되는 매크로블록 모드를 나타내는 방법.

청구항 34.

제32항에 있어서,

상기 라그랑지 곱수는 인터 프레임(Inter frame) 또는 인트라 프레임(Intra frame)에 대해 선택되는 방법.

청구항 35.

제33항에 있어서,

$$\lambda_B = f(\underline{QP}) \times \lambda_{I,P} \text{ 및}$$

$$f(\underline{QP}) = \max(2, \min(4, \frac{\underline{QP}}{6})) \text{ 인 방법.}$$

청구항 36.

제33항에 있어서,

$$\lambda_B = f(QP) \times \lambda_{Lp} \text{ 및}$$

$$f(QP) = \max(2, \min(4, \frac{QP+2}{8})) \text{인 방법.}$$

청구항 37.

인코더 로직(encoder logic)을 사용하여 비디오 이미지 데이터의 적어도 일부분을 인코딩하는 단계; 및

상기 인코더 로직이 상기 인코더 로직에 의해 채용되는 모션 벡터 예측 유형을 식별하는 구문 정보를 출력하도록 하는 단계를 포함하는 방법.

청구항 38.

제37항에 있어서,

상기 인코더 로직은 공간 모션 벡터 예측(spatial motion vector prediction) 및 시간 모션 벡터 예측(temporal motion vector prediction)을 포함하는 그룹으로부터 선택되는 적어도 하나의 모션 벡터 예측 유형을 선택적으로 채용하도록 구성되며, 레이트 디스토션 최적화(RDO) 관정에 적어도 부분적으로 기초하여 상기 구문 정보와 대응 인코딩된 비디오 이미지 데이터를 선택적으로 출력하도록 구성되는 방법.

청구항 39.

비디오 인코딩 관련 정보를 전달하는 데 사용되는 방법에 있어서,

비디오 데이터를 인코딩하는 단계; 및

상기 인코딩된 비디오 데이터의 구문 부분 내에 적어도 하나의 디스크립터(descriptor)를 선택적으로 설정하는 단계 -상기 디스크립터는 상기 비디오 데이터로 인코딩된 적어도 하나의 B 프레임에 관련되는 어댑티브 공간/시공간 인코딩(adaptive Spatial/Sptio-temporal encoding)을 식별함- 를 포함하는 방법.

청구항 40.

제39항에 있어서,

상기 비디오 데이터를 인코딩하는 단계는, JVT 표준에 따라 상기 비디오 데이터를 인코딩하는 단계를 포함하며, 상기 인코딩된 비디오 데이터의 구문 부분 내의 상기 적어도 하나의 디스크립터는 상기 인코딩된 비디오 데이터의 픽처 레이어 부분을 갖는 direct_mv_spatial 파라미터를 포함하는 방법.

청구항 41.

비디오 인코딩 관련 정보를 전달하는 데 사용되는 방법에 있어서,

비디오 데이터를 인코딩하는 단계; 및

상기 인코딩된 비디오 데이터의 구문 부분 내에 적어도 하나의 디스크립터를 선택적으로 설정하는 단계 -상기 디스크립터는 상기 비디오 데이터로 인코딩된 적어도 하나의 인터 프레임 내에 어댑티브 카피/모션 카피 스킵 모드(adaptive Copy/Motion-Copy skip mode)를 식별함- 를 포함하는 방법.

청구항 42.

제41항에 있어서,

상기 비디오 데이터를 인코딩하는 단계는, JVT 표준에 따라 상기 비디오 데이터를 인코딩하는 단계를 포함하며, 상기 인코딩된 비디오 데이터의 상기 구문 부분 내의 상기 적어도 하나의 디스크립터는 상기 인코딩된 비디오 데이터의 픽처 레이어 부분을 갖는 copy_mv_spatial 파라미터를 포함하는 방법.

청구항 43.

비디오 프레임의 정적 시간/공간(stationary temporal/spatial)을 고려한 비디오의 타임 스탬프 독립 모드 인코딩(time stamp independent mode encoding)에 사용되는 방법에 있어서,

비디오 시퀀스 내에 비디오 프레임의 적어도 일부분에 관련되는 모션의 공간 예측을 선택적으로 적용하는 단계; 및
다른 비디오 프레임의 참조 부분에 대한 시간 모션 예측 정보(spatial motion prediction information)가 0이면, 상기 모션의 공간 예측을 0으로 설정하는 단계를 포함하는 방법.

청구항 44.

제43항에 있어서,

상기 공간 예측은 공간 0 바이어스되는(spatial zero biased) 방법.

청구항 45.

제43항에 있어서,

상기 공간 예측은 상기 비디오 프레임 내에 다른 부분의 정적 서브 파티션(stationary subpartition)을 더 고려하며, 상기 서브 파티션들 중 적어도 하나가 거의 0인 대응 모션을 가지면, 상기 다른 부분도 0인 모션을 갖는 것으로 간주되는 방법.

청구항 46.

복수개의 서로 다른 양자화 함수를 사용하여 데이터의 적어도 일부분을 처리하여 복수개의 대응하여 양자화된 데이터 부분을 생성하는 단계; 및

적어도 하나의 임계치에 기초하여 상기 양자화된 데이터 부분들 중 하나를 선택적으로 출력하는 단계를 포함하는 동작을 이행하도록, 적어도 하나의 처리부를 구성하기 위한 컴퓨터 구현가능 명령을 갖는 컴퓨터 판독가능 매체.

청구항 47.

제46항에 있어서,

초기 데이터를 상기 데이터의 적어도 일부분을 포함하는 복수개의 부분으로 분할하는 단계; 및

적어도 하나의 분류 특성에 기초하여 상기 데이터의 적어도 일부분을 분류하는 단계를 더 포함하는 컴퓨터 판독가능 매체.

청구항 48.

제47항에 있어서,

상기 적어도 하나의 임계치는 상기 적어도 하나의 분류 특성에 관련된 컴퓨터 판독가능 매체.

청구항 49.

제47항에 있어서,

상기 초기 데이터는, 이미지 데이터, 비디오 데이터, 오디오 데이터, 및 스피치 데이터를 포함하는 그룹으로부터 선택되는 데이터를 포함하며, 상기 초기 데이터는 이미지 데이터 또는 비디오 데이터를 포함하며, 상기 적어도 일부분은 블록, 매크로블록, 슬릿, 슬라이스, 및 섹션을 포함하는 그룹으로부터 선택되는 컴퓨터 판독가능 매체.

청구항 50.

제47항에 있어서,

상기 적어도 하나의 분류 특성은, 예지 특성, 텍스처 특성, 평활 특성, 루미넌스 특성, 크로미넌스 특성, 컬러 특성, 노이즈 특성, 오브젝트 특성, 모션 특성, 사용자 선호도 특성, 사용자 인터페이스 포커스 특성, 레이아웃 특성, 타이밍 특성, 볼륨 특성, 주파수 특성, 피치 특성, 톤 특성, 품질 특성, 비트 레이트 특성, 데이터형 특성, 해상도 특성, 및 암호화 특성을 포함하는 특성들의 그룹으로부터 선택되는 컴퓨터 판독가능 매체.

청구항 51.

제46항에 있어서,

상기 복수개의 서로 다른 양자화 함수는 적어도 2개의 서로 달리 동작하는 데드존 양자화기를 포함하며, 상기 데드존 양자화기들 중 적어도 하나는 어댑티브 커버리지 데드존 양자화기를 포함하는 컴퓨터 판독가능 매체.

청구항 52.

제46항에 있어서,

상기 적어도 하나의 임계치에 기초하여 상기 양자화된 데이터 부분들 중 하나를 선택적으로 출력하는 단계는, 상기 양자화된 데이터 부분들 중 하나를 선택하기 위해서 레이트 디스토션 최적화(RDO)를 이행하는 단계를 더 포함하는 컴퓨터 판독가능 매체.

청구항 53.

이미지 데이터의 적어도 일부분에 대하여 적어도 하나의 특성 분석을 이행하는 단계;

상기 적어도 하나의 특성 분석에 기초하여 인코더 내의 적어도 하나의 어댑티브 양자화 파라미터를 선택적으로 설정하는 단계; 및

상기 인코더로 상기 이미지 데이터의 적어도 일부분을 인코딩하는 단계를 포함하는 동작을 이행하도록, 적어도 하나의 처리부를 구성하기 위한 컴퓨터 구현가능 명령을 갖는 컴퓨터 판독가능 매체.

청구항 54.

제53항에 있어서,

상기 적어도 하나의 어댑티브 양자화 파라미터는 상기 인코더 내에서 어댑티브 커버리지 데드존 양자화기에 관련되는 컴퓨터 판독가능 매체.

청구항 55.

제53항에 있어서,

상기 적어도 하나의 특성 분석에 기초하여 상기 인코더 내의 적어도 하나의 어댑티브 양자화 파라미터를 선택적으로 설정하는 단계는, 비디오 프레임을 포함하는 상기 이미지 데이터의 복수개의 부분 각각에 대하여 온 더 플라이 방식으로 이행되는 컴퓨터 판독가능 매체.

청구항 56.

적어도 2개의 서로 다른 데드존 양자화기를 사용하여 이미지 데이터의 적어도 일부분을 인코딩하는 단계; 및

적어도 하나의 판정 팩터에 관련되는 레이트 디스토션 최적화(RDO)에 기초하여 상기 적어도 서로 다른 데드존 양자화기들 중 하나의 출력에서 바람직한 인코드 데이터를 식별하는 단계를 포함하는 동작을 이행하도록, 적어도 하나의 처리부를 구성하기 위한 컴퓨터 구현가능 명령을 갖는 컴퓨터 판독가능 매체.

청구항 57.

제1 데드존 양자화기를 사용하여 이미지 데이터의 적어도 일부분을 인코딩하는 단계;

상기 제1 데드존 양자화기가 적어도 하나의 판정 팩터를 충족하는 지를 판정하는 단계;

만약 그렇다면, 상기 제1 데드존 양자화기의 출력을 출력시키는 단계;

그렇지 않다면, 상기 제1 데드존 양자화기와 다른 적어도 제2 데드존 양자화기를 사용하여 상기 이미지 데이터의 적어도 일부분을 인코딩하는 단계를 포함하는 동작을 이행하도록, 적어도 하나의 처리부를 구성하기 위한 컴퓨터 구현가능 명령을 갖는 컴퓨터 판독가능 매체.

청구항 58.

제57항에 있어서,

레이트 디스토션 최적화(RDO) 판정에 기초하여 상기 이미지 데이터의 적어도 일부분의 허용가능 인코드 버전을 식별하는 단계를 더 포함하는 컴퓨터 판독가능 매체.

청구항 59.

이미지 데이터의 적어도 일부분에 대하여 이미지 분석을 이행하는 단계;

상기 분석된 이미지 데이터 부분에 대하여 블록 분류를 이행하는 단계;

상기 블록 분류된 이미지 데이터 부분의 데드존 양자화를 이행하는 단계; 및

상기 데드존 양자화된 이미지 데이터 부분의 인코딩을 이행하는 단계를 포함하는 동작을 이행하도록, 적어도 하나의 처리부를 구성하기 위한 컴퓨터 구현가능 명령을 갖는 컴퓨터 판독가능 매체.

청구항 60.

제59항에 있어서,

상기 이미지 분석은 에지 탐지 및 텍스처 분석을 포함하는 그룹으로부터 선택되는 적어도 한 유형의 분석을 포함하는 컴퓨터 판독가능 매체.

청구항 61.

제59항에 있어서,

상기 블록 분류는 상기 이미지 분석에 기초하여 동작 구성되며, 상기 데드존 양자화는 상기 블록 분류에 기초하여 동작 구성되는 컴퓨터 판독가능 매체.

청구항 62.

적어도 2개의 서로 다른 인코더를 사용하여 비디오 이미지 데이터의 적어도 일부분을 인코딩하는 단계 -상기 2개의 서로 다른 인코더들 중 적어도 하나는 상기 비디오 이미지 데이터의 논 레지듀얼 모드를 지원하도록 동작 구성되는 데드존 양자화기를 포함함-; 및

적어도 하나의 판정 팩터에 관련되는 레이트 디스토션 최적화(RDO) 판정에 기초하여 상기 적어도 2개의 서로 다른 인코더의 출력 내에서 바람직한 인코드 프레임 데이터를 식별하는 단계를 포함하는 동작을 이행하도록, 적어도 하나의 처리부를 구성하기 위한 컴퓨터 구현가능 명령을 갖는 컴퓨터 판독가능 매체.

청구항 63.

제62항에 있어서,

상기 데드존 양자화기는 어댑티브하며, 상기 논 레지듀얼 모드가 사용하는 컴퓨터 판독가능 매체.

청구항 64.

제62항에 있어서,

상기 인코더들 중 적어도 하나는 상기 비디오 이미지 데이터의 다이렉트 모드를 지원하도록 동작 구성되는 컴퓨터 판독가능 매체.

청구항 65.

제62항에 있어서,

상기 인코더들 중 적어도 하나는 상기 비디오 이미지 데이터의 카피 모드를 지원하도록 동작 구성되는 컴퓨터 판독가능 매체.

청구항 66.

제62항에 있어서,

상기 논 레지듀얼 모드는 런 령쓰 인코딩(RLC) 방식에 기초하여 동작 구성되는 컴퓨터 판독가능 매체.

청구항 67.

제62항에 있어서,

상기 데드존 양자화기를 갖는 상기 인코더는, 크로미넌스 레지듀 데이터, 루미넌스 레지듀 데이터, 및 모든 레지듀 데이터를 포함하는 레지듀 데이터의 그룹으로부터 선택되는 적어도 한 유형의 레지듀를 제거하도록 동작 구성되는 컴퓨터 판독가능 매체.

청구항 68.

이미지 데이터의 적어도 일부분의 적어도 하나의 특성에 기초하여 양자화 함수를 갖는 인코딩 로직을 동작 구성하는 적어도 하나의 라그랑지 곱수를 선택적으로 변경하는 단계; 및

상기 인코딩 로직을 사용하여 상기 이미지 데이터의 상기 적어도 일부분을 인코딩하는 단계를 포함하는 동작을 이행하도록, 적어도 하나의 처리부를 구성하기 위한 컴퓨터 구현가능 명령을 갖는 컴퓨터 판독가능 매체.

청구항 69.

인코더 로직을 사용하여 비디오 이미지 데이터의 적어도 일부분을 인코딩하는 단계; 및

상기 인코더 로직이 상기 인코더 로직에 의해 채용되는 모션 벡터 예측 유형을 식별하는 구문 정보를 출력하도록 하는 단계를 포함하는 동작을 이행하도록, 적어도 하나의 처리부를 구성하기 위한 컴퓨터 구현가능 명령을 갖는 컴퓨터 판독가능 매체.

청구항 70.

비디오 데이터를 인코딩하는 단계; 및

상기 인코딩된 비디오 데이터의 구문 부분 내에 적어도 하나의 디스크립터를 선택적으로 설정하는 단계 -상기 디스크립터는 상기 비디오 데이터로 인코딩된 적어도 하나의 B 프레임과 관련되는 어댑티브 공간/시공간 인코딩을 식별함- 를 포함하는 동작을 이행하도록, 적어도 하나의 처리부를 구성하기 위한 컴퓨터 구현가능 명령을 갖는 컴퓨터 판독가능 매체.

청구항 71.

제70항에 있어서,

상기 비디오 데이터를 인코딩하는 단계는 JVT 표준에 따라 상기 비디오 데이터를 인코딩하는 단계를 포함하며, 상기 인코딩된 비디오 데이터의 상기 구문 부분 내에서 상기 적어도 하나의 디스크립터는 상기 인코딩된 비디오 데이터의 픽처 레이어 부분을 갖는 `direct_mv_spatial` 파라미터를 포함하는 컴퓨터 판독가능 매체.

청구항 72.

비디오 데이터를 인코딩하는 단계; 및

상기 인코딩된 비디오 데이터의 구문 부분 내에 적어도 하나의 디스크립터를 선택적으로 설정하는 단계 -상기 디스크립터는 상기 비디오 데이터로 인코딩된 적어도 하나의 인터 프레임 내에 어댑티브 카피/모션 카피 스킵 모드를 식별함- 를 포함하는 동작을 이행하도록, 적어도 하나의 처리부를 구성하기 위한 컴퓨터 구현가능 명령을 갖는 컴퓨터 판독가능 매체.

청구항 73.

제41항에 있어서,

상기 비디오 데이터를 인코딩하는 단계는 JVT 표준에 따라 상기 비디오 데이터를 인코딩하는 단계를 포함하며, 상기 인코딩된 비디오 데이터의 상기 구문 부분 내에서 적어도 하나의 디스크립터는 상기 인코딩된 비디오 데이터의 픽처 레이어 부분을 갖는 `copy_mv_spatial` 파라미터를 포함하는 컴퓨터 판독가능 매체.

청구항 74.

비디오 시퀀스 내에 비디오 프레임의 적어도 일부분에 관련되는 모션의 공간예측을 선택적으로 적용하는 단계; 및

다른 비디오 프레임의 참조 부분에 대한 시간 모션 예측정보가 0이면, 상기 모션의 공간예측을 0으로 설정하는 단계를 포함하는 동작을 이행하도록, 적어도 하나의 처리부를 구성하기 위한 컴퓨터 구현가능 명령을 갖는 컴퓨터 판독가능 매체.

청구항 75.

제43항에 있어서,

상기 공간예측은 공간 0 바이어스되는 컴퓨터 판독가능 매체.

청구항 76.

제43항에 있어서,

상기 공간예측은 상기 비디오 프레임 내에 다른 부분의 정적 서브 파티션을 더 고려하고, 상기 서브 파티션들 중 적어도 하나가 대략 0인 대응 모션을 가지면, 상기 다른 부분은 또한 0 모션을 갖는 것으로 간주되는 컴퓨터 판독가능 매체.

청구항 77.

복수개의 서로 다른 양자화 함수를 이용하여 데이터의 적어도 일부분을 처리하여 복수개의 대응하여 양자화된 데이터 부분을 생성하고,

적어도 하나의 임계치에 기초하여 상기 양자화된 데이터 부분들 중 하나를 선택적으로 출력하도록 동작 구성되는 로직을 포함하는 장치.

청구항 78.

제77항에 있어서,

상기 로직은, 초기 데이터를 상기 데이터의 적어도 일부분을 포함하는 복수개의 부분으로 분할하고, 적어도 하나의 분류 특성에 기초하여 상기 데이터의 적어도 일부분을 분류하도록 더 구성되며,

상기 적어도 하나의 임계치는 상기 적어도 하나의 분류 특성에 관련되는 장치.

청구항 79.

제77항에 있어서,

상기 복수개의 서로 다른 양자화 함수는 적어도 2개의 서로 달리 동작하는 데드존 양자화기를 포함하며,

상기 데드존 양자화기들 중 적어도 하나는 어댑티브 커버리지 데드존 양자화기를 포함하는 장치.

청구항 80.

제77항에 있어서,

상기 로직은 상기 양자화된 데이터 부분들 중 하나를 선택하기 위해서 레이트 디스토션 최적화(RDO)를 이행하도록 더 구성되는 장치.

청구항 81.

이미지 데이터의 적어도 일부분에 대하여 적어도 하나의 특성 분석을 이행하고, 상기 적어도 하나의 특성 분석에 기초하여 인코더 내에 적어도 하나의 어댑티브 양자화 파라미터를 선택적으로 확립하며, 상기 인코더 내에 상기 이미지 데이터의 적어도 일부분이 인코딩되도록 동작 구성되는 로직을 포함하는 장치.

청구항 82.

제81항에 있어서,

상기 적어도 하나의 어댑티브 양자화 파라미터는 상기 인코더 내에 어댑티브 커버리지 데드존 양자화기에 관련되는 장치.

청구항 83.

적어도 2개의 서로 다른 데드존 양자화기를 이용하여 이미지 데이터의 적어도 일부분이 인코딩되도록 하며, 적어도 하나의 판정 팩터에 관련된 레이트 디스토션 최적화(RDO)에 기초하여 적어도 2개의 서로 다른 데드존 양자화기들 중 하나의 출력 내에서 바람직한 인코딩 데이터를 식별하도록 동작하게 구성되는 로직을 포함하는 장치.

청구항 84.

제1 데드존 양자화기를 이용하여 이미지 데이터의 적어도 일부분이 인코딩되도록 하며, 상기 제1 데드존 양자화기의 출력이 적어도 하나의 판정 팩터를 충족하는지를 판단하여, 충족한다면, 상기 제1 데드존 양자화기의 출력을 출력시키고, 충족하지 않으면, 상기 제1 데드존 양자화기와 다른 적어도 제2 데드존 양자화기를 이용하여 상기 이미지 데이터의 적어도 일부분이 인코딩되도록 동작 구성되는 로직을 포함하는 장치.

청구항 85.

제84항에 있어서,

상기 로직은 레이트 디스토션 최적화(RDO) 판정에 기초하여 상기 이미지 데이터의 적어도 일부분의 허용가능 인코딩 버전을 식별하도록 더 구성되는 장치.

청구항 86.

상기 이미지 데이터의 적어도 일부분에 대한 이미지 분석을 이행하고, 상기 분석된 이미지 데이터 부분에 대하여 블록 분류를 이행하며, 상기 블록 분류된 이미지 데이터 부분의 데드존 양자화를 이행하고, 상기 데드존 양자화된 이미지 데이터 부분이 인코딩되도록 동작 구성되는 로직을 포함하는 장치.

청구항 87.

적어도 2개의 서로 다른 인코더를 이용하여 비디오 이미지 데이터의 적어도 일부분을 인코딩하도록 동작 구성되며, 상기 2개의 서로 다른 인코더들 중 적어도 하나는, 상기 비디오 이미지 데이터의 논 레지두얼 모드를 지원하며 적어도 하나의 판정 팩터에 관련된 레이트 디스토션 최적화(RDO) 판정에 기초하여 상기 적어도 2개의 서로 다른 인코더의 출력에서 바람직한 인코딩된 프레임 데이터를 식별하도록 동작 구성되는 데드존 양자화를 포함하는 장치.

청구항 88.

제87항에 있어서,

상기 데드존 양자화는 어댑티브하고 상기 논 레지두얼 모드를 선택적으로 사용하는 장치.

청구항 89.

제87항에 있어서,

상기 인코더들 중 적어도 하나는 상기 비디오 이미지 데이터의 다이렉트 모드와 상기 비디오 이미지 데이터의 카피 포드를 포함하는 그룹으로부터 선택되는 적어도 하나의 모드를 지원하도록 동작 구성되는 장치.

청구항 90.

이미지 데이터의 적어도 일부분의 적어도 하나의 특성에 기초하여 양자화 함수를 갖는 인코딩 로직을 동작 구성하는 적어도 하나의 라그랑지 곱수를 선택적으로 변경시키며, 상기 인코딩 로직을 이용하여 상기 이미지 데이터의 적어도 일부분이 인코딩되도록 동작 구성되는 로직을 포함하는 장치.

청구항 91.

인코더 로직을 이용하여 비디오 이미지 데이터의 적어도 일부분을 인코딩하며, 상기 인코더 로직이, 상기 인코더 로직에 의해 채용된 모션 벡터 예측 유형을 식별하는 구분 정보를 출력하도록 동작 구성되는 로직을 포함하는 장치.

청구항 92.

비디오 이미지를 인코딩하고, 상기 인코딩된 비디오 데이터의 구문 부분 내에 적어도 하나의 디스크립터를 선택적으로 설정하고, 상기 디스크립터는 상기 비디오 데이터로 인코딩된 적어도 하나의 B 프레임에 관련되는 어댑티브 공간/시공간 인코딩을 식별하도록 동작 구성되는 로직을 포함하는 장치.

청구항 93.

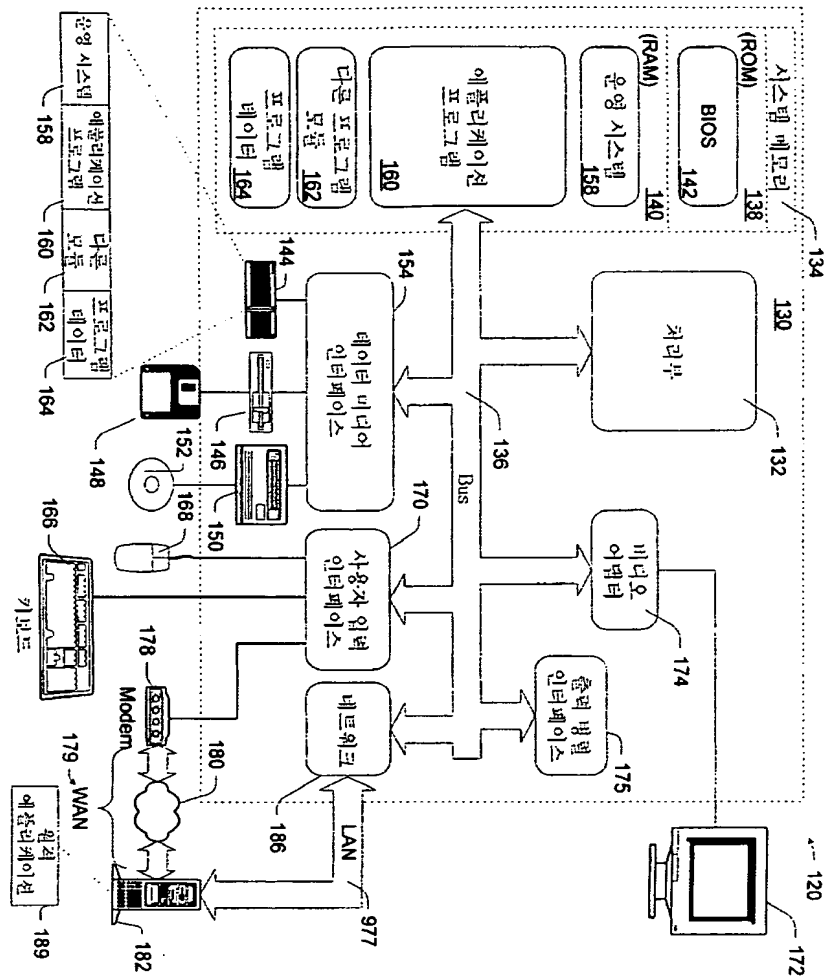
비디오 이미지를 인코딩하고, 상기 인코딩된 비디오 데이터의 구문 부분 내에 적어도 하나의 디스크립터를 선택적으로 설정하고, 상기 디스크립터는 상기 비디오 데이터로 인코딩된 적어도 하나의 인터 프레임에 관련되는 어댑티브 카피/모션 카피 스킵 모드를 식별하도록 동작 구성되는 로직을 포함하는 장치.

청구항 94.

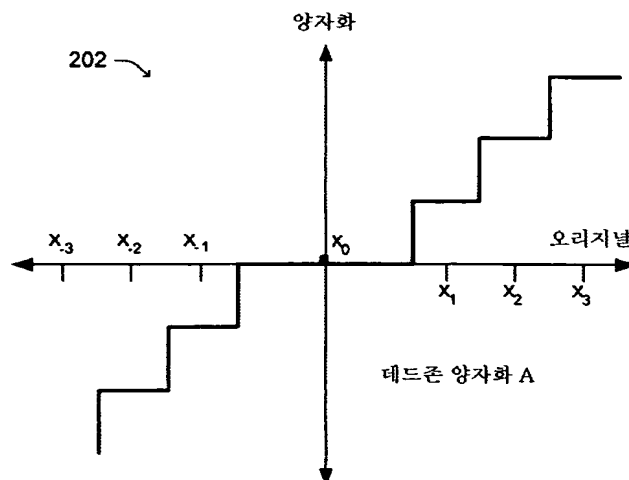
비디오 시퀀스 내에서 비디오 프레임의 적어도 일부분에 관련된 모션의 공간 예측을 선택적으로 적용하고, 만약 다른 비디오 프레임의 참조 부분에 대한 시간 모션 예측 정보가 0이면, 상기 모션의 공간 예측을 0으로 설정하도록 동작 구성되는 로직을 포함하는 장치.

도면

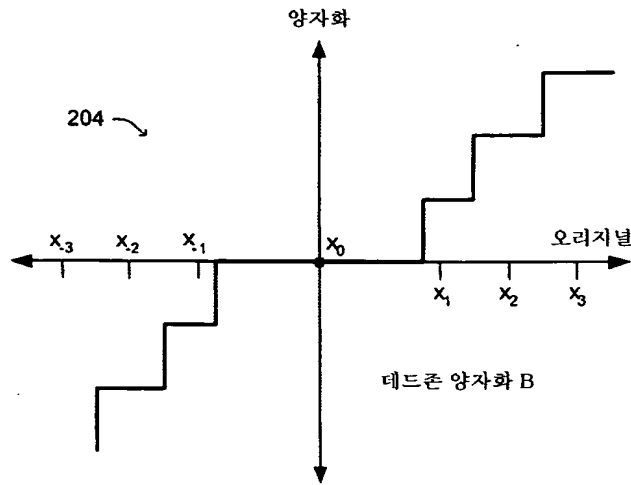
도면1



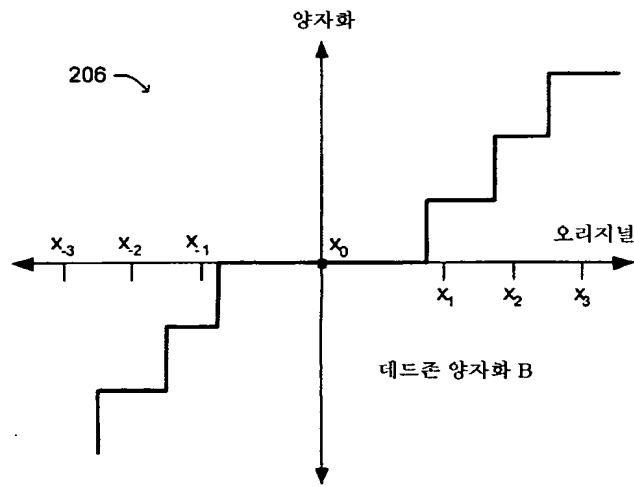
도면 2a



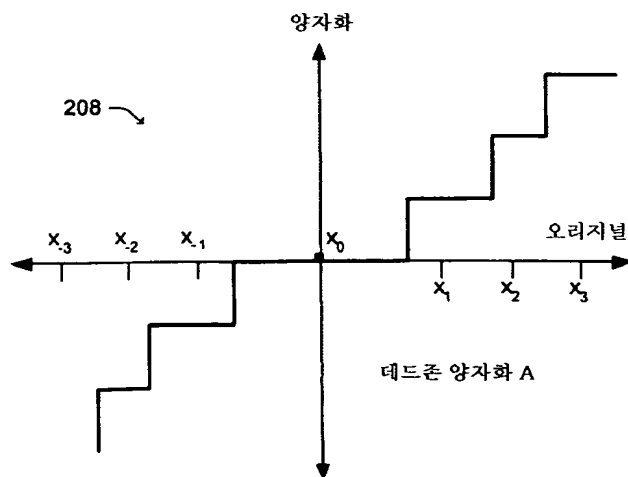
도면2b



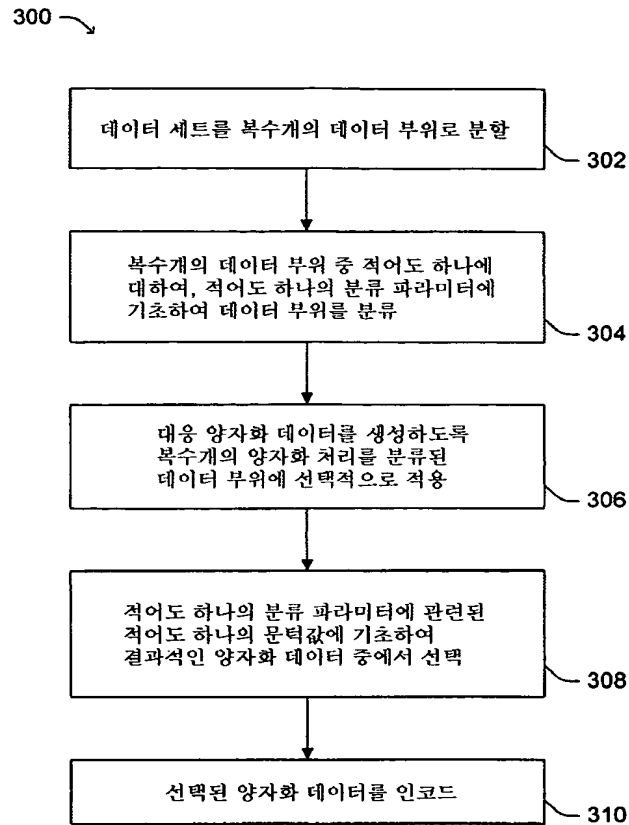
도면2c



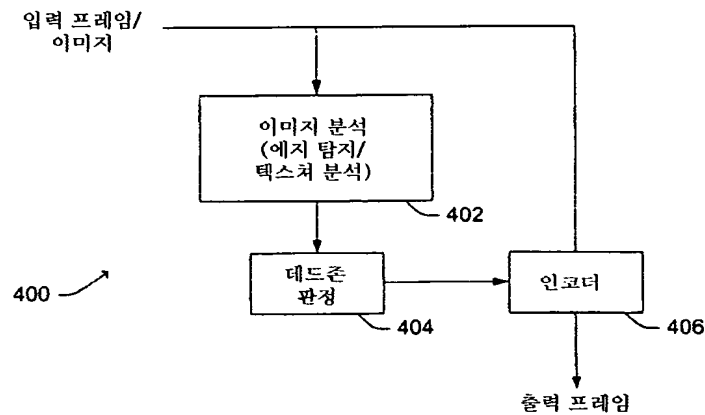
도면2d



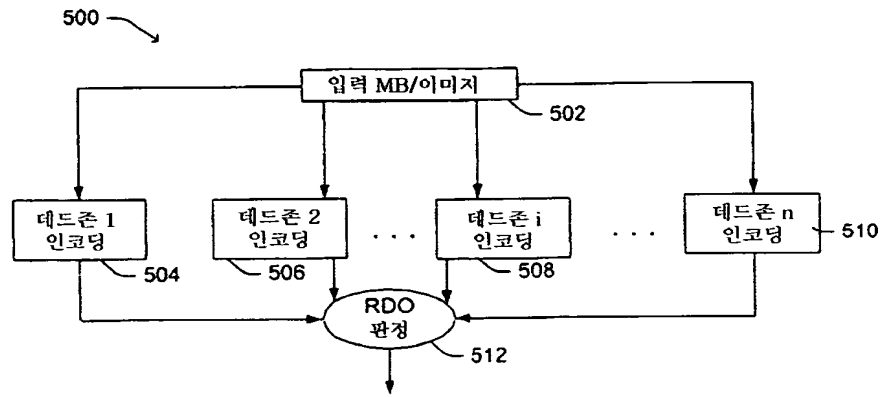
도면3



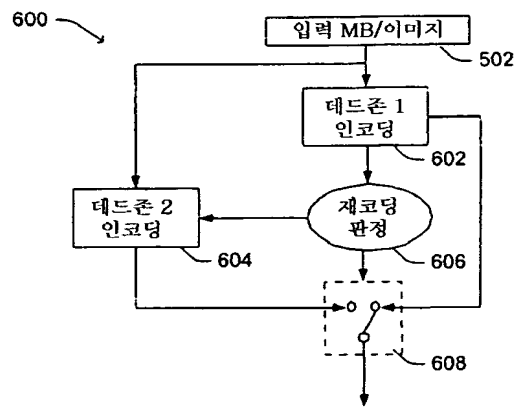
도면4



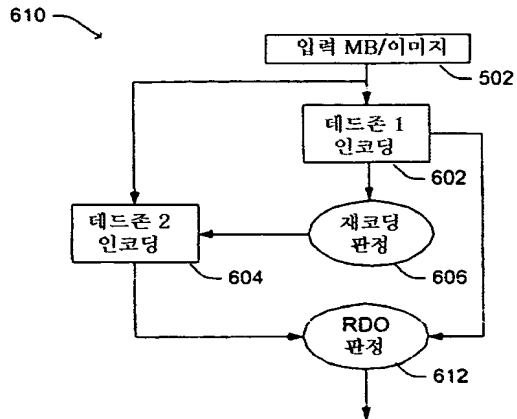
도면5



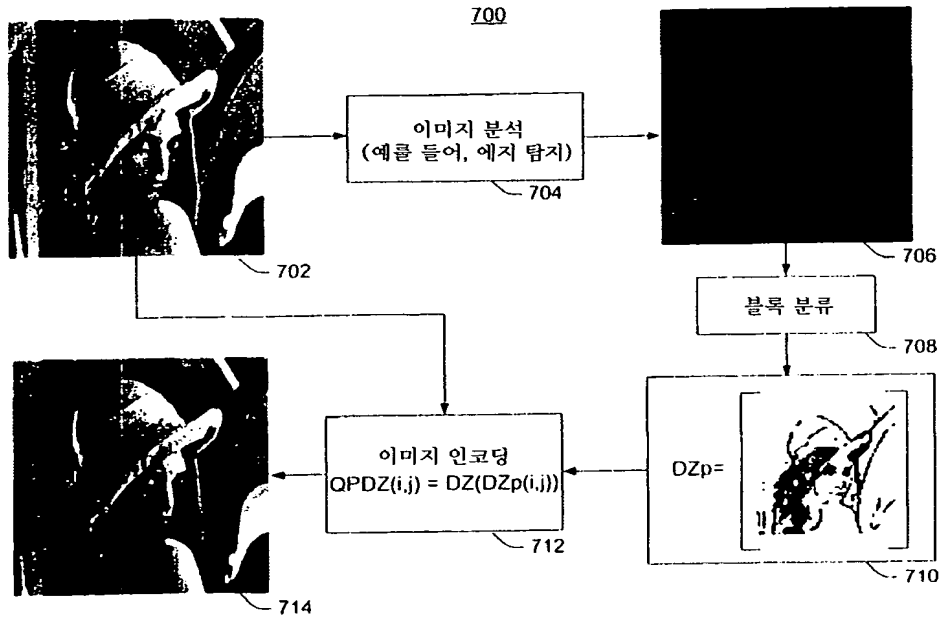
도면6a



도면6b



도면7

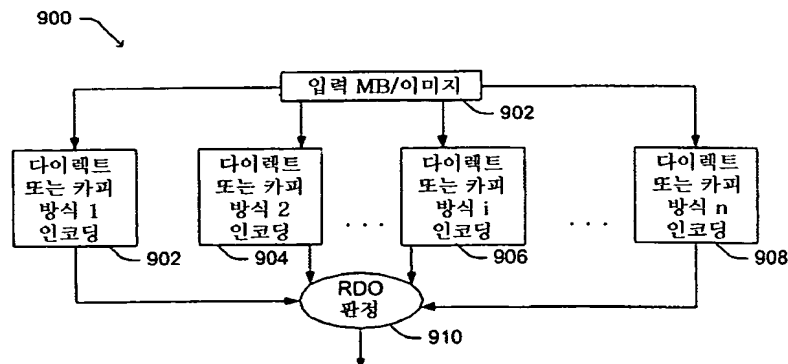


도면8

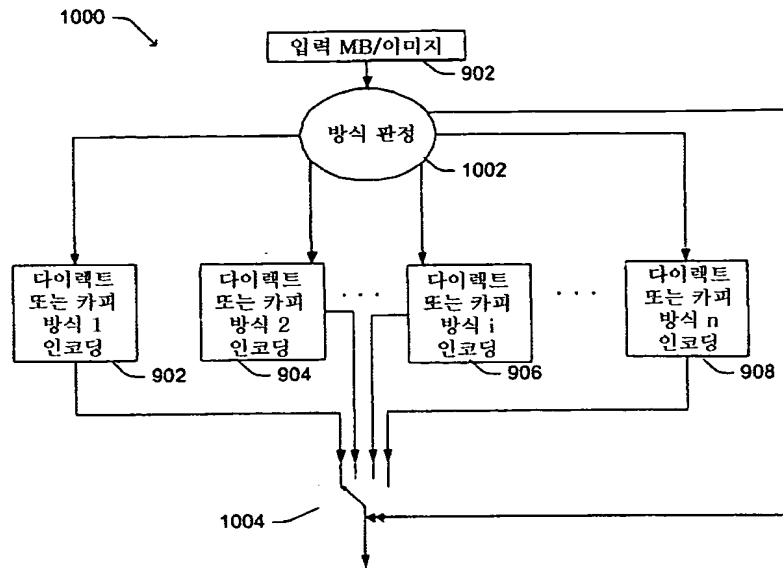
800

picture layer rbsp()	카테고리	디스크립터
picture structure	3	$e(v)$
frame_num	3	$u(v)$
rps layer()		
if(coding_type() == Inter) {		
copy_mv_spatial		$u(n)$ or $e(v)$
}		
if(coding_type() == Bipred) {		
direct_mv_spatial		$u(n)$ or $e(v)$
if(direct_mv_spatial) {		
direct_mv_scale_fwd	3	$e(v)$
direct_mv_scale_bwd	3	$e(v)$
direct_mv_scale_divisor	3	$e(v)$
}		
explicit_B_prediction_block_weight_indication		$e(v)$
if(explicit_B_prediction_block_weight_indication > 1)		
adaptive_B_prediction_coeff_table()		
}		
rbsp_trailing_bits()		
}		

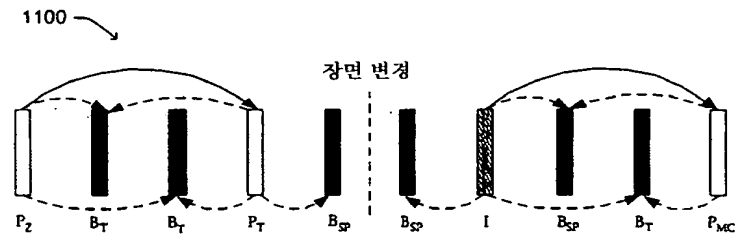
도면9



도면10



도면11



도면12

